User experience evaluation methods and models for software products

Yi Yang*, Feng Si, Jun Tan, Guoli Liu, Yunhong Zhang
National Key Laboratory of Human Factor and Ergonomics, China National Institute of Standardization, Beijing, 100191, China

ABSTRACT

Software can be improved through user experience metrics, so it is very important to choose the right measurement method. Through the desktop research and other research methods, the user experience research methods and model results are summarized. With software A and B as the research object, according to the product form and the characteristics of the business, the user experience evaluation model commonly used in the industry is used to measure the two software products by means of usability test and questionnaire data from the various attributes and indicators of the evaluation model, so as to the rationality and feasibility of the evaluation model and provide Suggestions for improvement as the research object.

Keywords: Software, model, metrics, user experience

1. INTRODUCTION

In recent years, user experience has attracted the attention of enterprises, and more enterprises have begun to adopt the user-centered design concept to design and measure user experience for products to improve product quality, enhance user satisfaction and repeat purchase rate, so as to win profits for enterprises. User experience can bring great value to the development of society, enterprises and users. Paying attention to user experience has gradually become one of the important goals of enterprise development.

The term "user experience" comes from the field of human-computer interaction[1], and also involves computer science, psychology, industrial design, communication, business management and other disciplines and professional fields[2]. At present, there is no unified definition of "user experience" in the industry. The definition of "user experience" in ISO 9241-210:2019 is widely accepted. The standard defines "user experience" as "people's perception of and response to a product, system, or service that is used or expected to be used".

The supplementary note to the definition refers to the scope of user experience as "all aspects of the user's feelings before, during, and after using a product or system, including emotions, beliefs, preferences, cognitive impressions, physical and psychological reactions, behaviours, and achievements". From the definition of user experience, it can be seen that user experience is dynamic. Users will experience early, middle and late stages in the process of using the product. Users have different feelings in different stages and in the overall process of using the product.

2. THE ROLE OF USER EXPERIENCE METRICS

In recent years, computer technology has made the concept of human-computer interaction deeply rooted in the hearts of the people. Now the evaluation of whether a software product is easy to use is not only measured from the perspective of system availability and system performance, but also pays more attention to the events that happen at each touch point when users use the product and the changes in users' subjective feelings caused by each event.

According to the definition of "user experience" in ISO 9241-210:2019, "user experience" is the total experience a user has before, during, and after using a product or system. Both short-term and long-term user experience will be affected by the time dimension, and the result of user experience will change. As a result, the measurement of user experience is very flexible. Enterprises should evaluate the software services provided to users comprehensively and measure the quality of service based on user behavior, so as to find and solve the problems in the process of product design and development. Management master Peter Drucker words: If you can't measure it, you can't improve it.

* yangyi@cnis.ac.cn

Third International Conference on Computer Science and Communication Technology (ICCSCT 2022)
edited by Yingfa Lu, Changbo Cheng, Proc. of SPIE Vol. 12506, 125062V
 · 0277-786X · doi: 10.1117/12.2662541

3. USER EXPERIENCE METRICS

Metrics is a method that can measure and evaluate products. It can obtain the status and changes of the product under test in a structured way. The measurement personnel can use these data to gain insight, determine the direction of optimization, and help the product improve. For a company, the process of strategically managing the overall user experience of a product or company service in order to promote the user-centered design concepts and principles of its products. Improve product quality and user satisfaction by measuring product users. As for designers, they can narrow the scope of problems by dividing scenes and quantitative design, and find the root cause of the existing problems of products.

A scientific and quantifiable user experience measurement system can focus on user experience problems for the purpose of determining the focus objects that can improve product experience, and solve user needs from the perspective of experience. At present, the international Organization for Standardization and enterprises generally agree that the user experience measurement methods are subjective measurement and objective measurement. Subjective measurement mainly adopts questionnaire survey, in-depth interview and other methods. These methods are highly subjective, so they rely on users' interpretation. Objective measurement methods collect users' physiological data or behaviors in real time with third-party measurement tools.

4. USER EXPERIENCE MEASUREMENT MODEL FOR SOFTWARE PRODUCTS

Nowadays, software products pay more and more attention to make qualified products according to standards and then continue to conduct rapid iteration to improve product quality. Software product quality model has been developing continuously for more than 40 years since it was proposed in 1970s.The McCall quality model and Boehm quality model came into being later, and the quality model contained in the model has been widely used. The latest ISO/IEC 25010 standard replaces ISO/IEC 9126 standard as the new software quality evaluation standard to make up for the deficiencies of ISO/IEC 9126 quality model in the application of emerging software and system evaluation [3-8].

The ISO/IEC 25010:2011 software Product quality model describes 8 quality characteristics and 36 quality sub-characteristics, as shown in Figure 1. The combination of these quality characteristics is called software quality model, and the characteristics are the reflection of software quality and can be used as evaluation criteria.

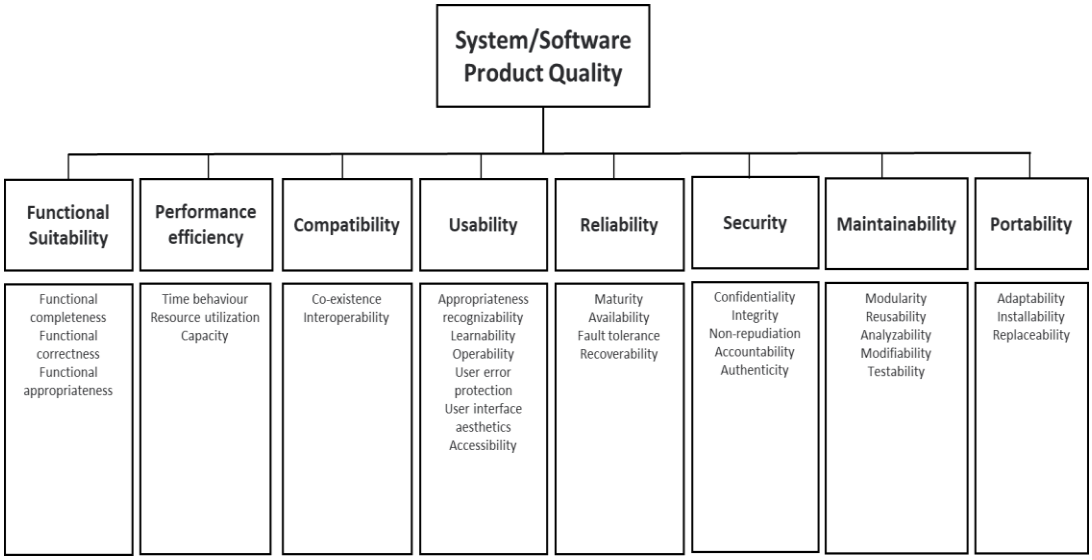

Figure 1. Product quality model.

ISO/IEC 25010:2011 Software quality is divided into use quality and product quality. Product quality refers to all inherent features of a product that can meet explicit and implicit requirements under specific use conditions. It reflects the degree to which a product meets product requirements, including functional suitability, performance efficiency, compatibility, usability, reliability, security, maintainability, and portability.

The use quality model is mainly composed of five quality characteristics, namely effectiveness, efficiency, satisfaction,

Freedom from risk and context coverage, as shown in Figure 2. Product usage quality is affected by product quality, hardware environment, operating system, user type, task type and social environment. In ISO/IEC 25010 standard, the quality of use is no longer a part of the quality of software products, but the embodiment of the quality of the whole system. The quality of use requires users to enjoy the process of using software products and trust the product. The product should not bring economic, health and environmental risks to users. Compared to the product model, the usage model focuses more on the role of the user of the product as the user. "Product quality model" and "Quality in use model" in ISO/IEC 25010 standard are widely used and famous quality models[9-13].

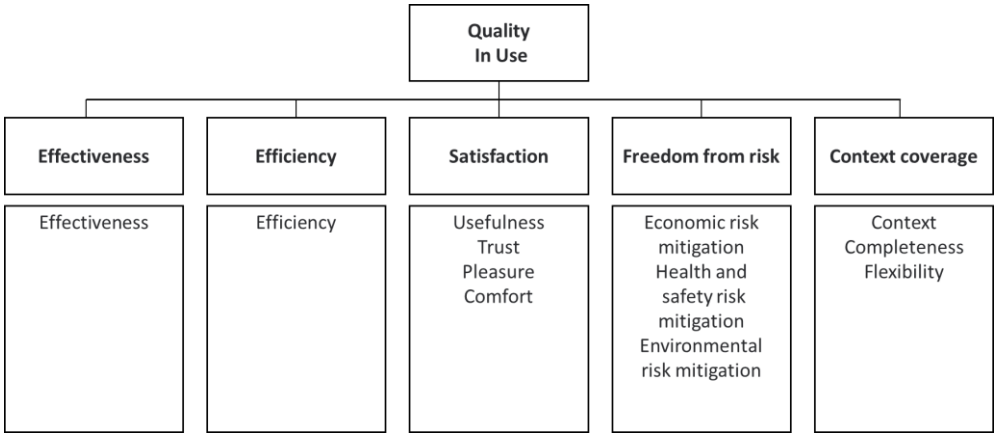

Figure 2. Quality in use model.

5. USER EXPERIENCE MEASUREMENT MODEL FOR SOFTWARE INDUSTRY

Software developers must develop products that meet the quality requirements according to the standardized software development standards. However, if the software only meets the requirements explicitly stated in the standards but does not meet other implicit requirements such as user expectations and personalized requirements, etc., the software quality may not reach the excellent level. Information technology industry are increasingly now, depending on the type of company and products for their products and system, designed a set of metrics for software quality comprehensive measures, focusing on business application end-to-end user experience[5], in order to meet user demand for software products more and look forward to. It is important and difficult to establish a user experience measurement model suitable for one's own business line. Before building the model, it is necessary to deepen the understanding of the existing experience measurement theories and measurement methods. Currently, mature experience measurement models are as follows [14-15]:

- The traditional Web product experience evaluation model "PLUSE" model proposed by Google, which includes Page Views, Uptime, Latency, Seven-day active users, Earnings;
- The "HEART" model is applied to user-centric web and mobile Internet networks, which includes Happiness, Engagement, Adoption, Retention and Task Success;
- The "PTECH" model suitable for enterprise-level middle and back office, which includes five indicators including Performance, Task Success, Engagement, Clarity and Happiness;
- Released by the ant gold suit in 2017 TECH model, this model includes the Task Success, Engagement, Clarity, Happiness five indexes such as Task completion rate;
- The "UES" model suitable for he cloud product usage experience measurement system, which includes five indicators such as Ease of use, Consistency, Happiness, Task Success, Performance.

After investigating the measurement model and software measurement system, this paper selects a new measurement model to test software products in a small range, aiming to output the performance scores of the products on various indicators and the problems exposed by the products in the measurement process to verify Measure the usefulness of the model. And refer to the measurement model, the motivation and background proposed in the desktop research, and the practical application scenarios, discuss the reliability of the measurement model and the selection of measurement

indicators that have been constructed, and provide a reference for the construction of software measurement models and software evaluation in the future.

6. THE RESEARCH METHODS

6.1 Objective

Based on the software A and B as the research object, according to the product form and the characteristics of the business, use the common user experience evaluation model from all attributes of the evaluation model and index, the two software product usability testing and questionnaire data combining qualitative and quantitative methods to measure and get the comprehensive evaluation, To further explore the rationality and feasibility of evaluation model and provide suggestions for improvement as the research object.

6.2 Measurement models and indicators

This article selects the classic software metrics to determine the availability of the model evaluation index for the Performance, Ease of use, Task Success, Consistency, and Happiness.

6.3 Participants

Users select ages 18 to 35, 10 people, average age was 24.56±5.65 years old. Among them, 5 are male and 5 are female. All users have experience in using similar products.

6.4 Instruments

Two laptop computer and a voice recorder and record the screen software.

6.5 Procedure

Test in laboratory simulation office environment, the environment was quiet, and the environmental illumination was constant. Testing requirements operating common two meetings of the software on the market, to complete the software installation and register, login name, booking meeting, enter the meeting, conference control and end the meeting. To balance users order to use the software. This test does not give users the opportunity to familiarize themselves with the task and the software. Users should familiarize themselves with the layout and content of the software interface in a fixed short period of time before the test.

The researcher used the usability test record form to record the user's task completion rate, overall satisfaction and other usability problems found by the user. After each sample task was completed, the researcher then interviewed all users and asked them to fill in the corresponding 10 point user experience questionnaire.

The researcher used a usability test record, record the user's task completion, overall satisfaction and other users find usability problems, and in each sample task was completed, then to user interviews of all users, and asked to fill in the corresponding scoring 10 user experience questionnaire.

6.6 Results

Staff will be sorting recycling effective questionnaire and input the data into Excel for numerical statistics, computing each site the average score of every index. There are 20 valid samples.

6.6.1 The total scores of metric. According to the user experience model software after the formula to calculate the overall user experience measure total score results for A = 4.58 software, software B = 4.21, as shown in Figure 3.

6.6.2 The scores of each metrics. Figure 3 shows that software A performs better than software B in performance indicators, and software A has a higher task completion rate than software B, one user for the first time when using the sample B "conference control" task if you don't be indirect tip is mission impossible. In terms of Consistency, software B has a slightly higher score than software A. In terms of ease of use, software A has a better performance than software B. In terms of all indicators, software A has a better performance than B and a higher user satisfaction.

The usability testing of the question to carry on the summary induction, the existing software A and B is mainly the following usability problems.

A: Software A's usability problems

- The function arrangement of level 3 of the main interface is not in line with users' habits.
- Some functional modules lack help information to guide users.
- Yet the adaptive screen, interface initial position on the right.

B: Software B's usability problems

- Interface layout is relatively complex, affect the initial users find the target function;
- "Voice transfer text" function can't achieve real-time display, delay longer;
- Individual function name font color shallow cause users can't find this feature in the interface position on it.

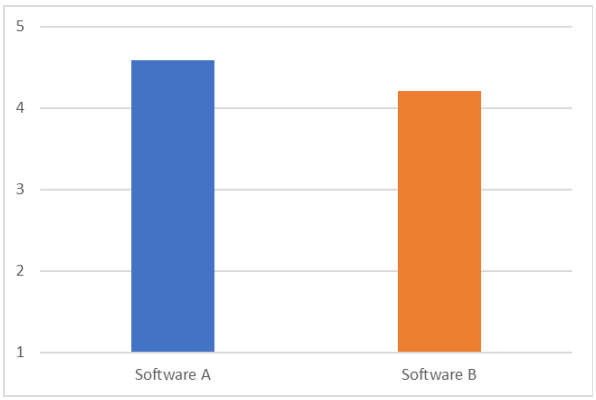

Figure 3. The total scores of metric.

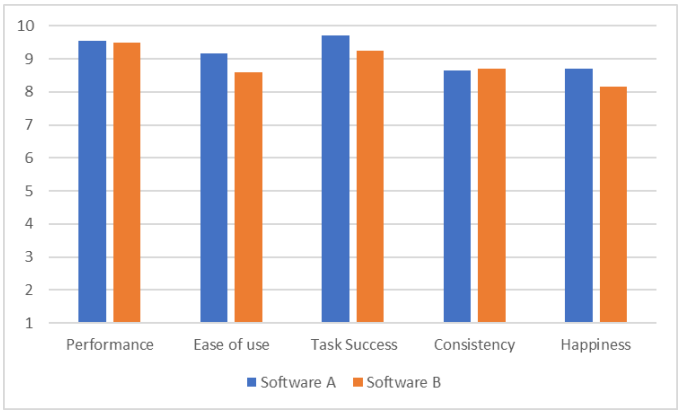

Figure 4. The scores of each metrics.

7. CONCLUSIONS

According to the user experience metrics, puts forward some improvement Suggestions about software A:

- Main interface of the third column with level 2 functions bar summary will merge user to more easily find the function and in line with the user's cognitive habits.
- If the user fails to operate a certain function, the interface should be real-time feedback, told that operation is not successful reason or solution is also given.
- The main interface is self-adaptive, and the screen is displayed in the center of the screen after the software is started.

For companies, software quality measurement models can be used to continuously improve quality, increase user stickiness and occupy more market share according to the characteristics of target user groups.

Software services for users, research can according to the measurement model objectively find out which factors will affect the users of the software product evaluation and feeling, and according to the measurements for the user to provide targeted policy recommendations. For a company, software quality measurement models can be used to continuously improve quality, increase user stickiness and occupy more market share according to the characteristics of target user groups.

8. RECOMMENDATIONS

In this measurement process, it is found that it is difficult to obtain the accuracy of some metrics. However, one metric can be divided into several measurable sub-indexes, and then the total score of the sub-indexes can be obtained by weighting the measurement scores of the sub-indexes. Therefore, in user experience evaluation, it is a major challenge to accurately measure how users feel when interacting with a product, system, or service. In the process of determining the measurement index, not all existing measurement indexes are measured, which will make the quality evaluation more complex, and even affect the objectivity of the evaluation. Quality features should be defined accurately, the relationship between quality features and indicators should be measured, which indicators are the key indicators to describe the user experience of the product to be measured, and quality features, sub-features, indicators should be assigned weights, and operational and quantifiable measurement methods should be developed to simplify complex problems.

ACKNOWLEDGMENT

This research was funded by Presidential Fund of CNIS (292022Y-9457-2).

REFERENCES

[1] Hu, F., and Jiang, M. Y., “Experience Design Research: Problem context, discipline logic and theoretical trends,” Packaging Engineering, 039(020), 60-75(2018).
[2] Zheng, Y., Zhu, Y., and Wang, H., “The development of user experience research, research models and evaluation methods,” Packaging Engineering, 41(6), 7(2020).
[3] Lima, L., Huve, C. A. and Peres, L. M., “Software product quality evaluation guide for electronic health record systems,” Proceedings of the 34th Brazilian Symposium on Software Engineering, 108-113(2020).
[4] Galli, T., Chiclana, F., and Siewe, F., “On the use of quality models to address distinct quality views,” Applied System Innovation, 4(3), 41(2021).
[5] Parthasarathy, S. and Sharma, S., “Impact of customization over software quality in ERP projects: An empirical study,” Software Quality Journal, 25(2), 581-598(2017).
[6] Wanigasooriya, C. S., Fernando, M. F. A. R., Madushanka, N. S. and Gunathilake, W., “Models, developments, trends, and evaluation of software product qualities: A review,” FOC Student Symposium, (2021).
[7] Zhao, H., and Chu, G. Y., “Study on application software usage quality based on ISO/IEC25010,” Value Engineering, 35(23), 2(2016).
[8] Shi, Y. J., and Zhou, R., “Construction and application of customer experience evaluation system,” Quality and Certification (12), 3(2018).
[9] Tamas, G., Francisco, C. and Francois, S, “Software Product Quality Models, Developments, Trends, and Evaluation,” SN Computer Science, 1(3), 154(2020).
[10] Argüello, M., Casanova, C. and Cedaro, K. E., “Proposal to improve software testing in small and medium enterprises,” Applied Informatics, 1455, 497-512(2021).
[11] Iqbal, J., Nasir, M. H. N., Khan, M., Awan, I. and Farid, S., “Software process improvement implementation issues in small and medium enterprises that develop healthcare applications,” Journal of Medical Imaging and Health Informatics, 10(10), 2393-2403(2020).
[12] Lenarduzzi, V., et al., “Open source software evaluation, selection, and adoption: A systematic literature review,” Euromicro Conference Series on Software Engineering and Advanced Applications, 437-444 (2020).
[13] Tsuda, N., et al., “WSQF: Comprehensive software quality evaluation framework and benchmark based on SQuaRE,” 2019 IEEE/ACM 41st International Conference on Software Engineering: Software Engineering in Practice (ICSE-SEIP) ACM, 312-321(2019).
[14] Alibaba Design, Produced by Ali! How do you quantify experience design? Just read this one!, https://www.uisdc.com/experience-design-metrics, (2021).
[15] PMCAFF, Superfull! Experience Metrics Theory 2021, https://blog.csdn.net/pmcaff2008/article/details/121240407, (2021).