



Usability testing of intensive course mobile application using the usability scale system

Manda Rohandi ^{a,1,*}; Nurlaila Husain ^{a,2}; Indri Wirahmi Bay ^{a,3}

^a Universitas Negeri Gorontalo, Jln. Jend. Sudirman, Kota Gorontalo 96128, Indonesia

¹ manda.rohandi@ung.ac.id; ² nurlailahusain@ung.ac.id; ³ indriwirahmi@ung.ac.id

*Corresponding author

Article history: Received March 30, 2021; Revised November 09, 2021; Accepted November 09, 2021; Available online December 07, 2021

Abstract

The Mobile Intensive Course (MIC) application version 2 is an application designed to assist the learning process of the English intensive course subject. Measurement of the application usability has never been done before, so the effectiveness, efficiency, and user satisfaction for this application have not been measured yet. Also, usability measurement can be used as a reference for the level of user loyalty, whether it is net promoters, passive users, or detractors. One of the usability measuring tools that are easy to use but still valid and reliable, is the System Usability Scale (SUS). The purpose of this study was to test the usability of the application to determine the quality using the SUS questionnaire. This research was conducted in four steps, namely 1) trying out the application to the respondents; 2) distributing SUS questionnaires; 3) calculating the average SUS score; 4) analyzing the average SUS score. This study involved 37 respondents consisting of students and EIC lecturers. The results of this study indicates that the usability of the application can be accepted by users with an average SUS score of 70.61 and obtained C mark based on the CGS assessment. Nevertheless, the application only obtained passive users. The average contribution value for learnability was still low, namely 1.9 from the maximum score 4. It is necessary to improve future applications in terms of application learnability that allowing users to be more familiar with applications and potentially becoming net promoters.

Keywords: Usability; System Usability Scale; MIC; EIC

Introduction

The Mobile Intensive Course (MIC) application is a website-base application for English learning [1], and is one type of Mobile-Assisted Language Learning application developed as an additional medium in learning English Intensive Course (EIC) courses [2], [3]. The material in the MIC application is taken from the EIC course textbook which consists of six units where each unit has four lessons. Each material contains four basic skills for listening, reading, writing and speaking. In addition to these four skills, this material also contains three components of English, vocabular, grammar, and pronunciation [4], [5]. The MIC application is a "do what I tell you" type of application, in which the user must follow the application instruction. This application consists of two users, lecturers and students.

MIC version 1 covered only two basic skills, listening and reading. Also, it does not provide video conferencing facilities. After developing the MIC version 2 application, it is then equipped with two basic skills of writing and speaking, as well as video conferencing facilities. According to the results of the evaluation of the MIC version 1 application design carried out previously in [2], the obtained results indicated that the EIC course material could assist students mastery, however the questions on the questionnaire covered more EIC materials than usability measurements (effectiveness, efficiency and effectiveness). Therefore, it is necessary to measure usability or the quality of user experience when using the application, so that users are expected to be more loyal to using the MIC version 2 application.

Usability is one of essential factors considered in designing interactive products [6]. According to ISO 9241-11 [7], usability is "the degree to which a product can be used by certain users to achieve certain goals through effectiveness, efficiency, and satisfaction, viewed from a particular context of use". A usable system must be easy to learn, efficient, easy to remember, and must also be able to meet user expectations and needs in the appearance and technical features of the application [8]. Usability is the main factor in developing and assessing the quality of a system or application seen from the user experience, and therefore the level of usability of an application is one of the factors that the application can be well received by users [9], [10], [11] [12].

There are several usability assessment questionnaires that can be used, such as 1) Questionnaire for User Interface Satisfaction (QUIS) which is a measurement tool to subjectively assess user satisfaction with special aspects of human interaction; 2) Software Usability Measurement Inventory (SUMI) which consists of 50 questions and used to measure perceptions, effectiveness, usability, system and user learnability of the system; 3) Post-Study Usability Questionnaires

(PSSUQ), used to measure user satisfaction with a product or system, this questionnaire has 16 question items [13], and 4) System Usability Scale (SUS) which is one most popular approach to measuring usability [14]. Although it is free, the SUS questionnaire is a valid and reliable test tool, so it is a measurement tool that is often used to measure usability [14], [15], [16]. In addition, SUS also has several other advantages, such as easy to use, does not require complex calculations, can be used even though the sample size is small [17], can be used for all types of systems [18], and has been tested on mobile applications [19] and also on the website.

Research on measuring the usability of a Tegal city government website using SUS has been carried out by [9]. The results obtained in this study indicated that the website is not yet usable, so it required improvement to be more effective, efficient and satisfying for its users. Other studies have also been conducted to reveal the usability differences of the four popular mobile applications, namely Facebook, WhatsApp, YouTube and E-Mail using the SUS measuring tool [8]. The findings of this study indicated that the four mobile applications had usability above the standard and very satisfying to their users, with WhatsApp obtained the highest score and Facebook obtained the lowest score. The two studies showed that the SUS measuring instrument was valid and reliable to measure user satisfaction in terms of the usability of an application

This study aims to test the usability of the system of MIC version 2 application through the SUS questionnaire. SUS was chosen as a tool to measure user experience, because SUS can use a small sample, the simple calculation, free of charge, and valid and reliable results. The results of the research can be used to help improve the MIC version 2 application to be more effective, efficient and more user-friendly.

Method

There are three types of methods in evaluating the usability of mobile applications, laboratory experiments, field studies, and direct measurements [20]. In this study, a laboratory experimental method was used because the respondents were in a controlled environment and performed specific tasks related to mobile applications [8].

The sample in this study was 37 people with 25 students registering EIC courses from 5 classes (consisting of male and female) and 12 EIC lecturers. According to [9], the feasible sample size in the study is between 30 to 500

This research was conducted through four steps; 1) MIC version 2 application was given to respondents to be installed into their mobile devices and then tested on respondents according to the instructions given; 2) The SUS questionnaire was distributed to respondents to be filled out online via google form and/or manually according to the respondent's experience after using the MIC version 2 application; 3) The results of the SUS questionnaire assessment by the respondents were then calculated to obtain the average SUS score; 4) Analyze the average SUS final score that has been obtained.

The SUS questionnaire has 10 basic and simple questions as shown in **Table 1**, and is a very useful tool in understanding the problems faced by users in using the system [8], [10], [12]. The structure of the SUS questionnaire is divided into two parts, namely "learnability" and "usability" [12], [21]. The questions on the questionnaire consist of two learnability questions (questions 4 and 10) and eight usability questions. The questionnaire asked respondents to fill ten agree or disagree questions (some positive and some negative) regarding the application being evaluated [22]. The rating scale on the SUS questionnaire uses a 5-point Likert scale, with a number 1 indicating "strongly disagree", number 2 "disagree", number 3 "neutral", number 4 "agree", and number 5 "strongly agree". If the respondent does not answer the question items on the questionnaire, then the scale value is filled with number 3 [1]. For question items with positive prompts (1, 3, 5, 7, and 9) the value of the contribution is obtained from the position of the scale minus 1, and for question items with negative prompts (2, 4, 6, 8, and 10) the contribution value is obtained by 5 deducted with the number of respondent's choice. The contribution value ranges from 0 to 4, and the final score is obtained from the total contribution value of each question item (maximum total contribution value = 40) multiplied by 2.5 to get a score of 100 [17].

Table 1. Questions of SUS

Question code	Question Items
Q1	I think that I would like to use this system.
Q2	I found the system unnecessarily complex.
Q3	I thought the system was easy to use
Q4	I think that I would need to support of a technical person to be able to use this system.
Q5	I found the various functions in the system were well integrated.
Q6	I thought there was too much inconsistency in the system.
Q7	I would imagine that most people would learn to use this system very quickly.
Q8	I found the system very cumbersome to use.
Q9	I felt very confident using the system.
Q10	I needed to learn a lot of things before I could get going with this system.

The following is the equation to calculate SUS score:

$$Skor\ SUS = ((Q1 - 1) + (5 - Q2) + (Q3 - 1) + (5 - Q4) + (Q5 - 1) + (5 - Q6) + (Q7 - 1) + (5 - Q8) + (Q9 - 1) + (5 - Q10)) * 2.5 \quad (1)$$

The SUS score ranges from 0 to 100, and an application can be said to have a good or acceptable *usability* category if the final SUS score is 68 [16] or 70 [17], and vice versa [9]. According to [24], [25], the SUS score scoring category uses the Curved Grading Scale (CGS), which is from A+ for the best score to F for the worst score, as shown in **Table 2**.

Table 2. Scoring guide of SUS according to CGS

SUS Score range	Marks	Percentage range
84.1 to 100	A+	96 to 100
80.8 to 84	A	90 to 95
78.9 to 80.7	A-	85 to 89
77.2 to 78.8	B+	80 to 84
74.1 to 77.1	B	70 to 79
72.6 to 74	B-	65 to 69
71.1 to 72.5	C+	60 to 64
65 to 71	C	41 to 59
62.7 to 64.9	C-	35 to 40
51.7 to 62.6	D	15 to 34
≤ 51.6	F	≤ 14

Results and Discussion

Responses were obtained from 25 students from 5 EIC random classes, plus 12 respondents from EIC lecturers. The results of the questionnaire were then calculated using equation 1 to obtain the SUS score and the average SUS score. The results of these calculations can be seen in **Table 3**.

Table 3. The result of SUS calculation

Respondent	Questions										SUS Score
	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q 10	
1	5	4	5	2	5	2	5	3	3	3	72,5
2	5	3	5	3	5	2	5	3	5	5	72,5
3	4	1	5	3	4	2	5	1	4	2	82,5
4	4	2	5	1	4	2	5	1	5	4	82,5
5	4	5	5	4	5	3	4	2	5	4	62,5
6	4	5	5	4	4	5	5	4	4	4	50
7	4	2	5	1	5	1	5	1	4	2	90
8	5	2	5	2	4	5	5	1	5	2	80
9	4	4	5	2	4	2	5	1	5	4	75
10	2	3	5	4	5	1	5	1	5	2	77,5
11	3	4	3	4	5	4	4	3	4	5	47,5
12	4	1	5	1	4	2	5	1	5	1	92,5
13	3	2	5	2	4	3	4	2	5	3	72,5
14	5	2	1	4	5	4	5	2	5	4	62,5
15	3	3	4	1	4	2	4	2	5	2	75
16	4	5	4	2	5	1	4	2	5	4	70
17	3	4	5	2	3	2	3	2	3	3	60
18	5	2	5	3	4	1	4	1	5	1	87,5
19	5	3	5	3	4	3	3	1	4	3	70
20	4	3	5	4	4	3	5	5	4	4	57,5
21	1	1	1	4	5	1	5	1	4	5	60
22	5	1	5	1	5	1	5	1	5	2	97,5
23	3	3	3	4	4	3	3	3	4	4	50
24	5	5	5	2	5	1	5	1	4	1	85
25	5	3	4	2	5	2	5	2	4	3	77,5
26	2	2	4	2	4	2	4	2	4	4	65
27	4	1	5	3	4	2	4	2	4	4	72,5
28	4	3	3	3	3	3	3	3	3	4	50
29	5	3	3	4	4	3	4	2	4	5	57,5
30	4	2	4	2	4	3	4	2	4	2	72,5

Respondent	Questions										SUS Score
	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	
31	5	4	4	4	4	3	5	2	4	5	60
32	4	2	4	2	4	2	4	2	4	4	70
33	5	2	5	4	4	2	4	2	4	2	75
34	5	2	4	4	4	2	4	1	4	5	67,5
35	5	1	5	1	5	1	5	1	1	5	80
36	5	2	4	4	4	2	4	2	3	4	65
37	5	2	4	4	4	2	4	2	4	4	67,5
	Rata-rata skor SUS										70,61

The contribution value of *usability* and *learnability* questions can be seen from the SUS questionnaire structure. It is found that the average contribution value for *usability* questions is 3.05 and *learnability* is 1.9 from the maximum contribution value of 4. The contribution value for each *usability* and *learnability* question can be seen in **Table 4**.

Table 4. Contribution score for *usability* and *learnability* questions on the SUS structure

Respondent	Contribution Score									
	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
1	4	1	4	3	4	3	4	2	2	2
2	4	2	4	2	4	3	4	2	4	0
3	3	4	4	2	3	3	4	4	3	3
4	3	3	4	4	3	3	4	4	4	1
5	3	0	4	1	4	2	3	3	4	1
6	3	0	4	1	3	0	4	1	3	1
7	3	3	4	4	4	4	4	4	3	3
8	4	3	4	3	3	0	4	4	4	3
9	3	1	4	3	3	3	4	4	4	1
10	1	2	4	1	4	4	4	4	4	3
11	2	1	2	1	4	1	3	2	3	0
12	3	4	4	4	3	3	4	4	4	4
13	2	3	4	3	3	2	3	3	4	2
14	4	3	0	1	4	1	4	3	4	1
15	2	2	3	4	3	3	3	3	4	3
16	3	0	3	3	4	4	3	3	4	1
17	2	1	4	3	2	3	2	3	2	2
18	4	3	4	2	3	4	3	4	4	4
19	4	2	4	2	3	2	2	4	3	2
20	3	2	4	1	3	2	4	0	3	1
21	0	4	0	1	4	4	4	4	3	0
22	4	4	4	4	4	4	4	4	4	3
23	2	2	2	1	3	2	2	2	3	1
24	4	0	4	3	4	4	4	4	3	4
25	4	2	3	3	4	3	4	3	3	2
26	1	3	3	3	3	3	3	3	3	1
27	3	4	4	2	3	3	3	3	3	1
28	3	2	2	2	2	2	2	2	2	1
29	4	2	2	1	3	2	3	3	3	0
30	3	3	3	3	3	2	3	3	3	3
31	4	1	3	1	3	2	4	3	3	0
32	3	3	3	3	3	3	3	3	3	1
33	4	3	4	1	3	3	3	3	3	3
34	4	3	3	1	3	3	3	4	3	0
35	4	4	4	4	4	4	4	4	0	0

Respondent	Contribution Score									
	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10
36	4	3	3	1	3	3	3	3	2	1
37	4	3	3	1	3	3	3	3	3	1
average /questions	3,1	2,3	3,3	2,2	3,3	2,7	3,4	3,1	3,2	1,6
Average <i>usability</i> contribution value = (total average value of question contribution 1,2,3,5,6,7,8, and 9)/8 = $(3,1+2,3+3,3+3,3+2,7+3,4+3,1+3,2)/8 = 3,05$ The average value of the learnability contribution = (total average value of the contribution of questions 4 and 10)/2 = $(2,2+1,6)/2 = 1,9$										

SUS is a global assessment of a system of a *usability* (effectiveness, efficiency and satisfaction) given by system users subjectively. From the results of the calculation of the SUS score in table III, the average SUS score of each question answered by all respondents is 70.61. These results indicate that the level of user acceptance of the MIC application is acceptable or has good usability, as shown in **Figures 1 and 2**. However according to the CGS assessment category in table II, the application received C mark as shown in **Figure 3**.

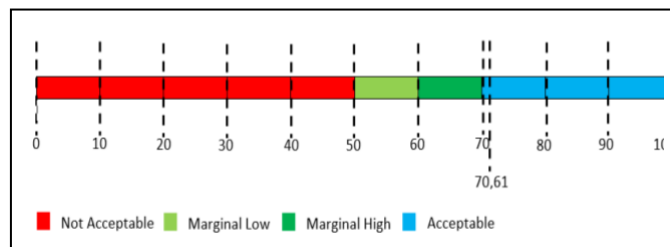


Figure 1. Acceptance rate of MIC applications according to [17]

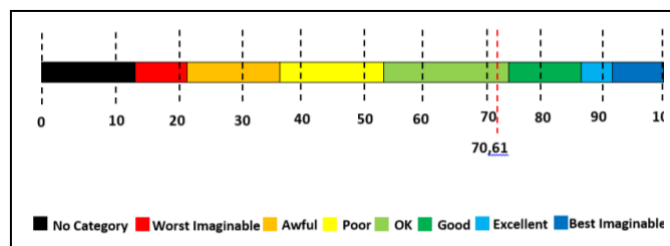


Figure 2. Acceptance rate of MIC applications according to [26]

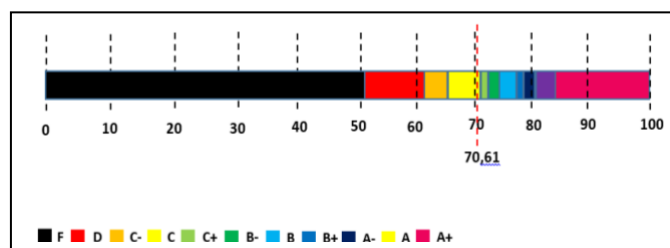


Figure 3. Acceptance rate of MIC applications according to [24], [25]

The results of the SUS score correspond to the average contribution value for the learnability question = 1.9 and *usability*= 3.05 (contribution value 0-4) on the SUS structure. According to the average learnability contribution value obtained, it shows that the developed MIC version 2 application still requires improvement so that it can be easier to use whereas the value of the *usability* contribution incites that the MIC version 2 application is considered quite good. Furthermore, the average contribution value of each positive and negative question item shows that the average contribution value of positive questions was greater than negative question items.

The SUS score can measure both the *usability* of the system and user loyalty, whether they are users who have the potential to become net promoters or detractors [27]. The average score of SUS if it has a value of ≥ 82 indicates the user has the potential to become a promoter, and if it has a value ≤ 67 then the user has the potential to become a detractor. Meanwhile, if the SUS score has a value of > 67 or < 82 , then it has the potential to become a passive user.

The SUS score for the MIC version 2 application, which is worth 70.61, shows that the loyalty of users of the MIC version 2 application only has the potential to be passive users who do not become promoters or detractors.

In general, the *usability* of the MIC version 2 application according to the results of the SUS calculation has a C mark or it is accepted by users according to CGS. Due to the low average value of the learnability contribution, the user is only a potential to be passive users. Therefore, improvements are needed in terms of simplifying the appearance of the features and functions of the application.

Conclusion

The quality of a system or application can be seen from the *usability* measurement perceived by the user. Because the high and low *usability* of an application is one of the factors inciting the application acceptability. One of the measuring tools that is often used in *usability* testing is the SUS questionnaire. This is because it is a valid and reliable *usability* testing tool. The results of *usability* test on the MIC version 2 application using the SUS questionnaire show that the application's SUS score is 70.61. This indicates that the MIC version 2 application is quite effective, efficient and satisfying for the user, so that it is accepted by the user and gets a C mark based on the CGS assessment category. This result is in accordance with the average contribution value for the *usability* and *learnability* question structure obtained with a value of 3.05 and 1.9 respectively from the maximum contribution value of 4. The level of user loyalty according to the SUS score obtained shows application users is only potentially passive users. From these results, it is necessary to improve the MIC version 2 application in order to provide a higher *usability* value. The *usability* improvement is in terms of memory efficiency because the application runs on mobile devices that have limited resources, as well as improving the effectiveness of the application in terms of the higher accuracy and the application improvement in carrying out the given task, so that users feel satisfied and give a positive response to the MIC version 2 application. *Usability* testing by using SUS can detect general *usability* issues and provide general solutions as well [8]. Therefore, comprehensive testing is needed to diagnose problems using scales, checklists or other *usability* tests to provide a deeper view and understanding of the *usability* of MIC version 2 applications.

References

- [1] M. Rohandi, N. Husain and I. W. Bay, "Development of Web-Based Computer-Assisted Language Learning in English Intensive Course," in *Annual Applied Science and Engineering Conference (AASEC)*, Bandung, 2017.
- [2] M. Rohandi, N. Husain and I. W. Bay, "Mobile-Assisted Language Learning Application for English Intensive Course," in *International Conference on Technology and Vocational Teachers (ICTVT)*, Yogyakarta, 2017.
- [3] M. Rohandi, N. Husain and I. W. Bay, "Pengembangan Mobile-Assisted Language Learning Menggunakan User Centered Design," *Jurnal Nasional Teknik Elektro dan Teknologi Informasi*, vol. 7, no. 1, pp. 27-33, Februari 2018.
- [4] N. R. Usu, Z. Tanipu, R. Kadir, Y. Bantulu, F. Mohammad, M. S. Ibrahim, N. Husain, I. W. Bay, M. E. W. Kau, S. W. Ali, H. Badu, F. Malabar and J. Saud, *Intensive Course of English: Students Book 1*, 1 ed., Gorontalo: UNG Press, 2019.
- [5] N. R. Usu, Z. Tanipu, R. Kadir, Y. Bantulu, F. Mohammad, M. S. Ibrahim, N. Husain, I. W. Bay, M. E. W. Kau, S. W. Ali, H. Badu, F. Malabar and J. Saud, *Intensive Course of English: Students Book 2*, 2 ed., Gorontalo: UNG Press, 2019.
- [6] Y. Rogers, Sharp and J. Preece, *Interaction Design: Beyond Human-Computer Interaction*, 3 ed., New York: John Wiley and Son, Ltd, 2002.
- [7] 9.-1. ISO, "Ergonomic Requirements for Office Work With Visual Display Terminals (VDTs), Part II Guidance of Usability," 1998.
- [8] A. Kaya, R. Azturk and C. A. Gummusoy, "*Usability*Measurement of Mobile Application with System *Usability*Scale (SUS)," in *Industrial Engineering in The Big data Era*, Cham, 2019.
- [9] I. Aprilia H. N., P. I. Santoso and R. Ferdiana, "Pengujian *Usability*Website Menggunakan System *Usability*Scale," *Jurnal Ilmu Pengetahuan dan Teknologi Komunikasi (IPTEK-KOM)*, vol. 17, no. 1, pp. 31-38, Juni 2015.
- [10] H. Anam, M. Sadiq and H. Jamil, "Development of System *Usability*Scale for the Urdu Language," *International Journal of Computer Science and Information Security (IJCSIS)*, vol. 18, no. 6, pp. 73-78, 2020.
- [11] Usability.gov, "*Usability*Evaluation Basics," 2020. [Online]. Available: <http://usability.gov/what-and-why/usability-evaluation.html>. [Accessed 20 Agustus 2020].
- [12] Z. Sharfina and H. B. Santoso, "An Indonesian Adaptation of the System *Usability*Scale (SUS)," in *ICACISIS*, 2016.
- [13] A. Garcia, "UX Research: Standardized *Usability*Questionnaire," 2013. [Online]. Available: <http://chaione.com/ux-reseach-standardizing-usability-questionnaire/>. [Accessed 15 6 2020].

-
- [14] Drew, R. Mandy, B. Falcone and W. L. Baccus, "What Does The System *Usability*Scale (SUS) Measure?," in *International Conference in Learning of Design, User Experience, and Usability*, Cham, 2018.
- [15] J. Xiong, "SUSapp: A Free Mobile Application That Makes the System *Usability*Scale (SUS) Easier to Administer," *Journal of UsabilityStudies*, vol. 15, no. 3, pp. 135-144, 2020.
- [16] J. Sauro, "Measuring *Usability*with The System Usability," 2011. [Online]. Available: <https://measuringu.com/sus/>. [Accessed 15 June 2020].
- [17] J. Brooke, "SUS: A Retrospective," *Journal of UsabilityStudies*, vol. 8, no. 2, pp. 29-40, 2013.
- [18] O. Konstantina, T. Nikolaos and K. Kristos, "Perceived *Usability*Evaluation of Learning Management Systems: Empirical Evaluation of The System *Usability*Scale," *The International Review of Research in Open and Distributed Learning*, vol. 16, no. 2, pp. 227-246, 2015.
- [19] P. Kortum and Sober, "Measuring the *Usability*of Mobile Application for Phones and Tablets," *International Journal of Human-Computer Interaction*, vol. 31, no. 8, pp. 518-529, 2015.
- [20] F. Nayebi, J. M. Desharnais and A. Abran, "The State of the Art of Mobile Application *Usability*Evaluation," in *25Th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, 2012.
- [21] Y. Wang, T. Lei and X. Liu, "Chinese System *Usability*Scale: Translation, Revision, Psychological Measurement ,," *International Journal of Human-Computer Interaction*, 2019.
- [22] S. McLellan, A. Muddimer and S. C. Peres, "The Effect Experience On System *Usability*Scale Ratings Vol. 7, No. 2, Februari 2012, Hal. 56-67," *Journal of UsabilityStudies*, vol. 7, no. 2, pp. 56-67, February 2012.
- [23] J. R. Lewis, "The System *Usability*Scale: Past, Present, and the Future," *International Journal of Human-Computer Interaction*, 2018.
- [24] J. Sauro and J. R. Lewis, *Quantifying the User Experience: Practical Statistics for User Research*, 1 ed., Cambridge: Morgan Kaufmann, 2012.
- [25] J. Sauro and J. R. Lewis, *Quantifying the User Experience: Practical Statistics for User Research*, 2 ed., Waltham: Morgan Kaufmann, 2016.
- [26] A. Bangor, P. Kortum and J. Miller, "Determining What Individual SUS Score Mean: Adding and Adjective Rating Scale," *Journal of UsabilityStudies*, vol. 4, no. 3, pp. 114-123, 2009.
- [27] J. Sauro, "Does Better *Usability*Increase Customer Loyalty?," 9 September 2010. [Online]. Available: <https://measuringu.com/usability-loyalty/>. [Accessed 2020].