Evaluation of Higher Education Learning Management System Usability Using System Usability Scale

Atika Herawati

(atikaherawati_1212821040@mhs.unj.ac.id)

Universitas Negeri Jakarta

Muchlas Suseno

(muchlas-suseno@unj.ac.id)

Universitas Negeri Jakarta

Siti Drivoka Sulistyaningrum

(drivoka@unj.ac.id)

Universitas Negeri Jakarta

During pandemic COVID-19, higher education has made a large shift in turning synchronous learning into asynchronous learning through online learning platforms. It encourages every institution, likewise Universitas Negeri Jakarta, to provide an all-in-one technology-supported learning environment due to face this dynamic change. Therefore, UNJ has finally developed its own Learning Management System (LMS). However, since it is a new brand tool of learning and innovation, it needs to be quickly evaluated for satisfaction and optimization so that there can be proposed for improvements in the future. This research uses System Usability Scale (SUS) as the Usability evaluation methods which is distributed through Google Forms and measured the validity by SPSS software. This evaluation method has been done to 50 participants from undergraduate and graduate students of English Language Education Study Program at UNJ. It was found that the result was valid with 0,702 reliability rate. The evaluation result of LMS owned by UNJ has 83,25 SUS score which can be regarded as "Excellent" rating as well as grade “A” according to Curved Grade Scale (CGS). Despite the score of Net Promoter Scale categorized as “Promoter”, it suggests further research investigated in other departments of UNJ to meet different expectations.

Keywords: Evaluation, Learning Management System (LMS), System Usability Scale (SUS), Higher Education

INTRODUCTION

The pandemic of coronavirus has caused disruptions in the teaching and learning process in higher education institutions. As a result of the outbreak, institutions were forced to conduct all of their activities with students completely online (Sobaih et al., 2020). Along with that, many governments took precautions to prevent the virus from spreading and to maintain the continuity of the educational process, and institutions all over the world adopted online learning (Ali, 2020). While Samir Abou El-Seoud et al. (2014) regarded online learning as an alternative to traditional learning, the coronavirus pandemic forced the growth of digital literacy and became an essential element for schools and higher education to maintain activity.

The higher education system is in a continuous process of change, and universities have to keep pace with the needs, desires, and requirements of students. Thus, information technologies and e-learning systems are seen as essential factors in carrying out the activities of universities; these institutions invest more and more in online systems and devices (Coman et al., 2020). This encourages every educational institution, including universities, to compete in optimizing online learning models through the development of various online learning (e-learning) platforms and to shift from synchronous to asynchronous learning with a Learning Management System (LMS). Historically, learning management systems (LMS) have become increasingly attractive in the past few years. Fearnley & Amora (2020) found that technical innovations have redefined the teaching and learning process. It all started in 1924 with the invention of the teaching machine (p.91).

Considering the aforementioned factors, this kind of online learning may be considered a logical evolution of the notion of remote learning in this epidemic era (Coman et al., 2020). A more sophisticated and encompassing definition argues that online learning with LMS is a type of teaching and learning that incorporates electronic resources and media with the goal of fostering development and improving education and training quality (p.3). Furthermore, an LMS is considered a system used for formal education or a network where knowledge is disseminated to a broad audience via electronic resources. Computers and the internet are the primary components that ensure the operation of such systems (p. 3). However, in this day and age, with the influence of the COVID-19 epidemic, one of the primary issues for universities is the usability of novel learning management systems.

The basis of this research is based on the observation that, until now, the learning management system owned by Universitas Negeri Jakarta (UNJ) has mostly been utilized as a medium for lecturers to provide learning materials and students to submit assignments in the form of text-based documents. Because it is new, other features and menus, such as attendance, quizzes, or multimedia-based assignments, are rarely designed to support more asynchronous learning models. As a result, Leavitt (2006) and Supriyadi et al. (2020) explained how the usability of the LMS is determined by the usability criteria, which must match user expectations about navigation, content, and website structure. Usability relates to the degree to which a product can be utilized by certain users to achieve specified goals effectively, efficiently, and satisfactorily in a usage environment (Kaya, Ozturk, & Gumussoy, 2020). Usability testing is

required to improve the user experience. Furthermore, usability is used to assess the amount of user satisfaction with a product, which is an LMS (Borsci et al., 2015).

Finally, this research aims to evaluate the usability of the UNJ LMS in undergraduate and graduate level English Language Education Study Programs. These findings may then be used as the foundation for continual improvement by UNJ LMS developers in order to deliver an online learning system that is more effective and efficient while also meeting the requirements and expectations of users, particularly students. The System Usability Scale is one of the tools used to assess a product's usability (SUS). Brooke created SUS as a "quick and dirty" usability evaluation instrument (2013). This helps inventors quickly evaluate their own products before mass distribution. According to Grier et al. (2013) and Lewis (2018), SUS is the most commonly used usability testing instrument for measuring usability perception (perceived usability) and is considered valid and reliable (Sauro & Lewis, 2011).

METHODOLOGY

This study employs quantitative research methodologies, including quantitative data analysis on the evaluation of learning management system usability using the System Usability Scale (SUS). This stage of research is depicted in Figure 1 and progresses iteratively from analysis through problem creation to drawing conclusions.

Problems Analysis and Formulation Stage

The stages of problem analysis and formulation must be highlighted in order to identify the research background. This stage is carried out by observing problems in research objects and also finding literature studies relevant to the problem to determine the scope and problem-solving method.

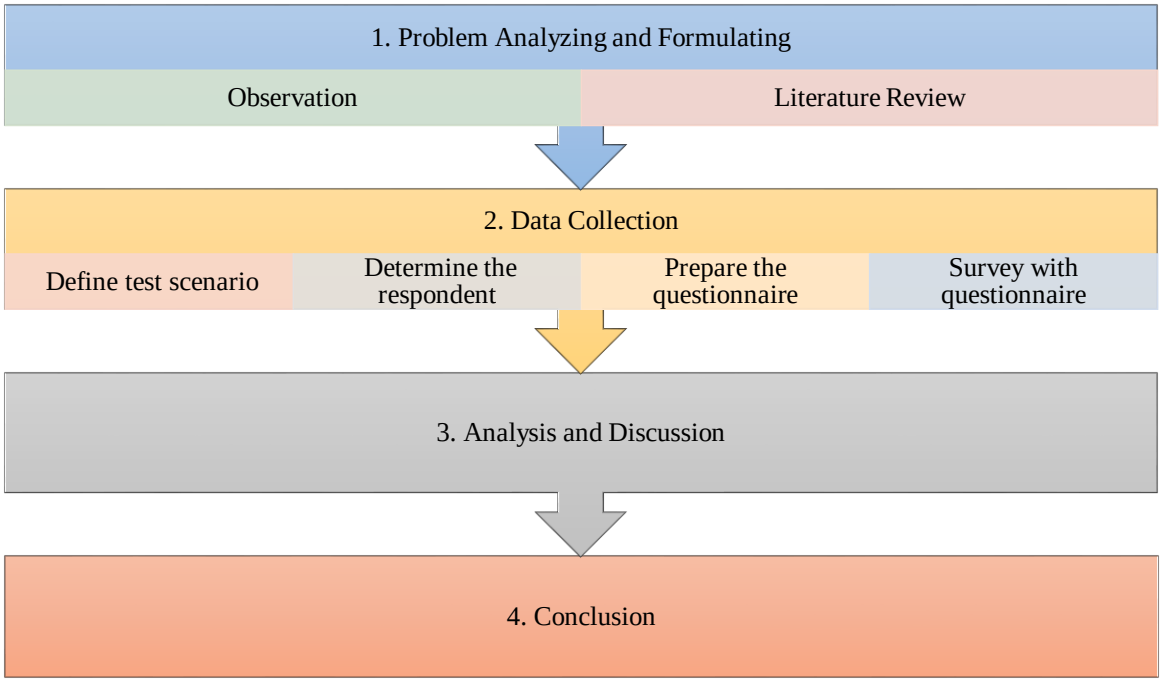

Figure 1. Research Stage

Data Collection Stage

The stage of data collection is separated into various sections. The initial step is to select the test scenario by discussing the UNJ LMS that will be tested and filling out the questionnaire. The next stage is the selection of responders, which will include 50 undergraduate and graduate students from the English Language Education Study Program who will utilize the LMS. Each undergraduate and graduate program required 25 individuals. This is acceptable since the study's sample size was between 30 to 500 people, according to Sugiyono's research (2008). The next step is to create questionnaires. The System Usability Scale (SUS) questionnaire is used to assess the usability of the UNJ LMS from the user's subjective perspective. SUS was chosen for this study because it has various advantages, including ease of use and usability outcomes in the form of scores ranging from 0 to 100 (Brooke, 2013), the analytic calculation procedure is simple (Bangor et al., 2009), and SUS demonstrated valid and reliable. The SUS questionnaire, according to Brooke (2013), consists of ten indicator statements, as indicated in Table 1. The assessment scale was based on a Likert scale of 1 to 5, with 1 indicating strong disagreement and 5 indicating strong agreement with each of the SUS statement components (Ependi & Panjaitan, 2018). The questionnaire was created in the form of a Google form and disseminated for two weeks, from October 1st to October 14th, 2021.

Code	Item of Statement	Scale
Q1	*I think that I would like to use the UNJ LMS frequently.*	1 to 5

Q2	*I found the UNJ LMS unnecessarily complex.*	1 to 5
Q3	*I thought the UNJ LMS was easy to use.*	1 to 5
Q4	*I think that I would need the support of a technical person to be able to use the UNJ LMS.*	1 to 5
Q5	*I found the various functions in the UNJ LMS were well integrated.*	1 to 5
Q6	*I thought there was too much inconsistency in the UNJ LMS.*	1 to 5
Q7	*I would imagine that most people would learn to use the UNJ LMS very quickly*	1 to 5
Q8	*I found the UNJ LMS very cumbersome (awkward to use.*	1 to 5
Q9	*I felt very confident using the UNJ LMS.*	1 to 5
Q10	*I needed to learn a lot of things before I could get going with the UNJ LMS*	1 to 5

Table 1. System Usability Scale (SUS) Item of Statement

Analysis and Discussion Stage

This stage is completed in order to compute the outcomes of the UNJ LMS evaluation using the SUS framework. At this stage, there is a rule that states that each statement item with an odd number (P1, P3, P5, P7, and P9) reduces the respondent's assessment score by 1, while each statement item with an even number (P2, P4, P6, P8, and P10) reduces the respondent's assessment score by 5. Then the next number of scoring scores is multiplied by 2.5. (Ika et al., 2015). The formula for calculating individual SUS scores is drawn as follows:

$$Score\ of\ SUS = (((Q1 - 1) + (5 - Q2) + (Q3 - 1) + (5 - Q4) + (Q5 - 1) + (5 - Q6) + (Q7 - 1) + (5 - Q8) + (Q9 - 1) + (5 - Q10)) \times 2{,}5)$$

The individual's SUS score is determined using as many as a number of respondents, and the average overall scoring score of each SUS evaluation of the individual is then calculated. This level of analysis and discussion also tested the statistical validity and reliability of respondent data by using SPSS software.

Conclusion Stage

After the average overall evaluation score of SUS and the validity and reliability test results are known, the next stage is to draw conclusions based on the results of usability research from the UNJ LMS.

FINDINGS AND DISCUSSION

Conclusion Stage

The data consisted of 50 SUS questionnaire results from UNJ LMS users. The profile of responders to the UNJ LMS is shown in Table 2. According to the distribution, respondents were undergraduate and graduate students from UNJ's English Language Education Study Program. According to that, the respondents consisted of 50 participants, with a proportion of 50% undergraduate students and 50% graduate students.

Scope	Frequency	Percentage
S1/Undergraduate	25	50%
S2/Graduate	25	50%
TOTAL	100	100%

Table 2. Respondents Profile and Distribution

Table 3 shows an example of data obtained from the UNJ LMS usability questionnaire using the SUS architecture. The table data is made up of respondent columns worth 1–50 respondents, with Q1 to Q10 as the codes for each statement in the SUS framework's ten statement items. Meanwhile, the score column is the column where the assessment of each respondent against each statement item is calculated using a formula (1). The value in the score column was also calculated using the formula (1), as can be seen below:

$$Score\ of\ SUS = \Big(\big((5-1)+(5-2)+(4-1)+(5-2)+(4-1)+(5-1)+(5-1) + (5-1)+(4-1)+(5-Q1)\big)\times 2{,}5\Big) = 87{,}5$$

The score is based on responses from people who have evaluated the UNJ LMS. Following the calculation of 50 respondents' data into the assessment score, the next stage is to compute the average SUS assessment score of 50 respondents, which is 83,25.

Respondents	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	SUS
1	5	2	4	2	4	1	5	1	4	1	87,5
2	4	1	4	2	4	2	5	2	5	1	85
3	4	1	4	1	3	1	4	1	4	1	85
4	4	1	4	1	4	2	4	3	5	2	80
…	…	…	…	…	…	…	…	…	…	…	…
…	…	…	…	…	…	…	…	…	…	…	…
50	4	1	5	1	3	1	3	1	3	3	77,5
Average score of SUS											**83,25**

Table 3. Data Generated from the SUS Questionnaire

Result of Validity Test

A validity test was performed to verify the validity or acceptability of questionnaires used to measure and collect research data from respondents. This validity test was carried out using SPSS software, with a correlation of coefficients using Pearson (2 tail) and a significance value of 5%. (Supriyadi et al., 2020). According to the data from 50 respondents, the value of R_{table} is 0,279. The validity test result of this study showed that the suitability of the questionnaire used is valid because $R_{count} > R_{table}$, that the score of R_{count} from each item is above 0,279 (R_{table}), as shown in table 4.

Item	R_{count}	R_{table}	Description
Q1	0,674	0,279	Valid
Q2	0,463	0,279	Valid
Q3	0,661	0,279	Valid
Q4	0,678	0,279	Valid
Q5	0,740	0,279	Valid
Q6	0,675	0,279	Valid
Q7	0,821	0,279	Valid
Q8	0,607	0,279	Valid
Q9	0,707	0,279	Valid
Q10	0,479	0,279	Valid

Table 4. Result of Statistical Validity Test

Result of Reliability Test

A reliability test was performed to see whether the questionnaire had consistency or reliability when measurements were repeated using the questionnaire. Aligned with Safitri et al. (2020), a questionnaire instrument is regarded as reliable if it achieves the following criteria:

Correlation Coefficient	Reliability Criteria
$0,81 < r \leq 1,00$	Very High
$0,61 < r \leq 0,80$	High
$0,41 < r \leq 0,60$	Enough
$0,21 < r \leq 0,40$	Low
$0,00 < r \leq 0,20$	Very Low

Table 5. Interpretation of Reliability

Cronbach's Alpha from the data gathered reveals a value of 0.702 based on statistical reliability test results performed using SPSS software, as shown in Table 6. This test questionnaire has a high level of reliability, as displayed in Table 5.

Cronbach's Alpha	N of Items
.702	10

Table 6. Result of Statistical Reliability Test

SUS Score Analysis

Upon obtaining the average score of the assessment of 50 respondents, the following stage is to determine the acceptance score grade using either of two methods (Brooke, 2013). The first method shows the acceptability ranges starting from "Not Acceptable"; "Marginal Low"; "Marginal High"; to "Acceptable". The second method is to rate the adjective rating, which consists of "Worst Imaginable"; "Poor"; "OK"; "GOOD"; "Excellent"; and "Best Imaginable". The average SUS score is considered to represent the value or level of user acceptance of the system tested for usability. To determine if the system has a usability level within the "Acceptable" range, the score of SUS must be greater than 70 (Brooke, 2013). As shown in Figure 2, the SUS score from the UNJ LMS is 83,25.

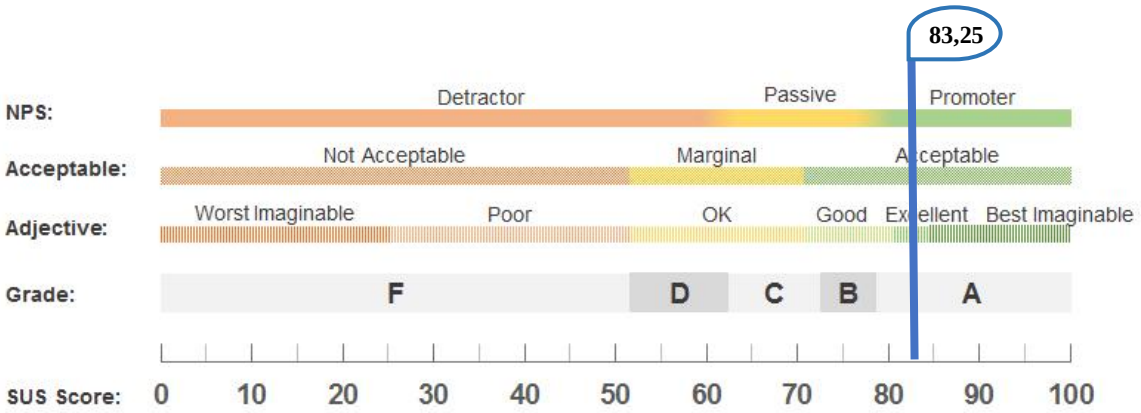

Figure 2. The UNJ LMS Acceptance Score Grade

The finding implies that the system is still in the range of achieving "Acceptable" and "Excellent" adjectival ratings. Furthermore, when using these two acceptance score grade techniques, the SUS scores can be interpreted in the Curved Grading Scale (CGS) evaluation form (Lewis, 2018). According to the grade scale, the UNJ LMS SUS score of 83,25 falls into the A category, as indicated in Table 7.

Grade	SUS	Percentile Range
A+	84.1 – 100	96 – 100
A	80.8 – 84.0	90 – 95
A-	78.9 – 80.7	85 – 89
B+	77.2 – 78.8	80 – 84
B	74.1 – 77.1	70 – 79
B-	72.6 – 74.0	65 – 69
C+	71.1 – 72.5	60 – 64
C	71.1 – 72.5	60 – 64
C-	62.7 – 64.9	35 – 40
D	51.7 – 62.6	15 – 34
F	0 – 51.6	0 – 14

Table 7. Curved Grade Scale (CGS) SUS adopted from J. R. Lewis and J. Sauro (2018)

Lewis (2018) has explained that the 10 components of the SUS statement are divided into two categories. Statements 1, 3, 5, 7, and 9 are positive statement items, while statements 2, 4, 6, 8, and 10 are negative statement items. As the data is acquired, the proportion and average of each category of positive and negative remarks are displayed in Table 8, Figure 3, and Figure 4. According to the statistics, the average responder ranked positive statement items higher than negative statement items (4,18 and 1,52, respectively). The results below show that respondents had a higher positive assessment of the UNJ LMS.

Category	**Item**	**Score**					**Average**	
		1	**2**	**3**	**4**	**5**		
Positive Items	Q1	0%	0%	2%	70%	28%	4,26	4,18
	Q3	0%	0%	2%	74%	24%	4,22	
	Q5	0%	0%	12%	52%	36%	4,24	
	Q7	0%	0%	16%	66%	18%	4,02	
	Q9	0%	0%	6%	72%	22%	4,16	
Negative Items	Q2	78%	20%	2%	0%	0%	1,24	1,52
	Q4	52%	40%	8%	0%	0%	1,56	
	Q6	66%	32%	2%	0%	0%	1,36	
	Q8	36%	56%	8%	0%	0%	1,72	
	Q10	48%	30%	22%	0%	0%	1,74	

Table 8. Result of Percentage and Average Score of Questionnaire

Table 8 shows the evaluation result in the use of the UNJ LMS as follows:

a) The average respondent stated that the UNJ LMS was easily used and that they could solve problems when first using the system.
b) The average respondent stated that the UNJ LMS was effective and efficient for them as well as the resources used to concern the results (how quickly) they achieved in completing a particular task.
c) The average respondent stated that the UNJ LMS had memorability where they used the system after not using it for a certain time, and they could still use it without difficulty.
d) The average respondent stated that the UNJ LMS found a minimum error when they used the system and it was easy to recover from this error.
e) The average respondent stated that the UNJ LMS satisfied them enough in terms of cognitive and emotional responses resulting from the use of a system, product, or service that met their needs and expectations.

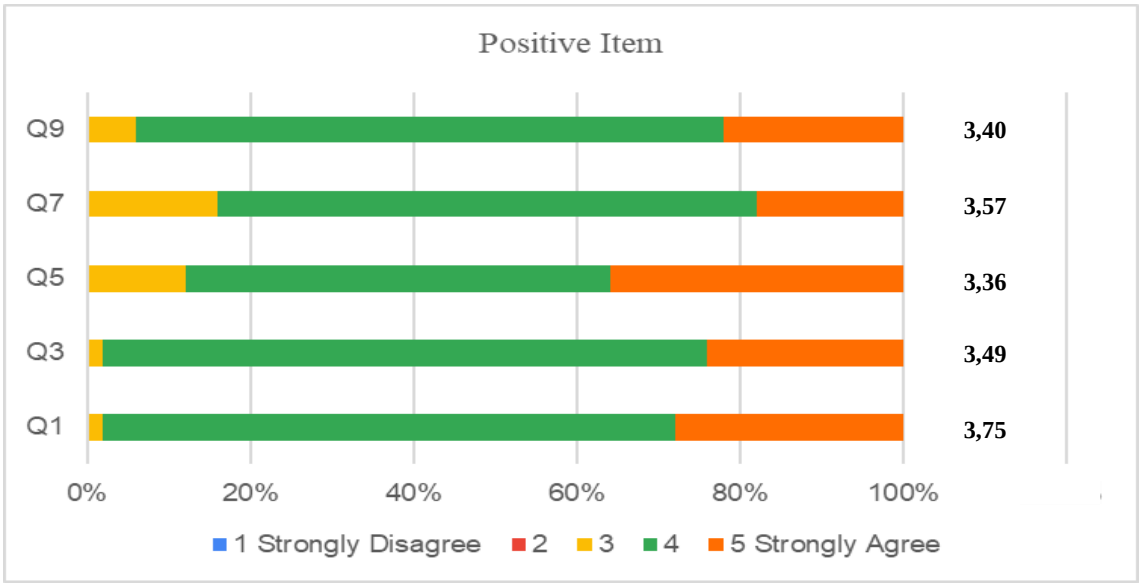

Figure 3. Percentage Spread and Average Value of Positive Statement Items

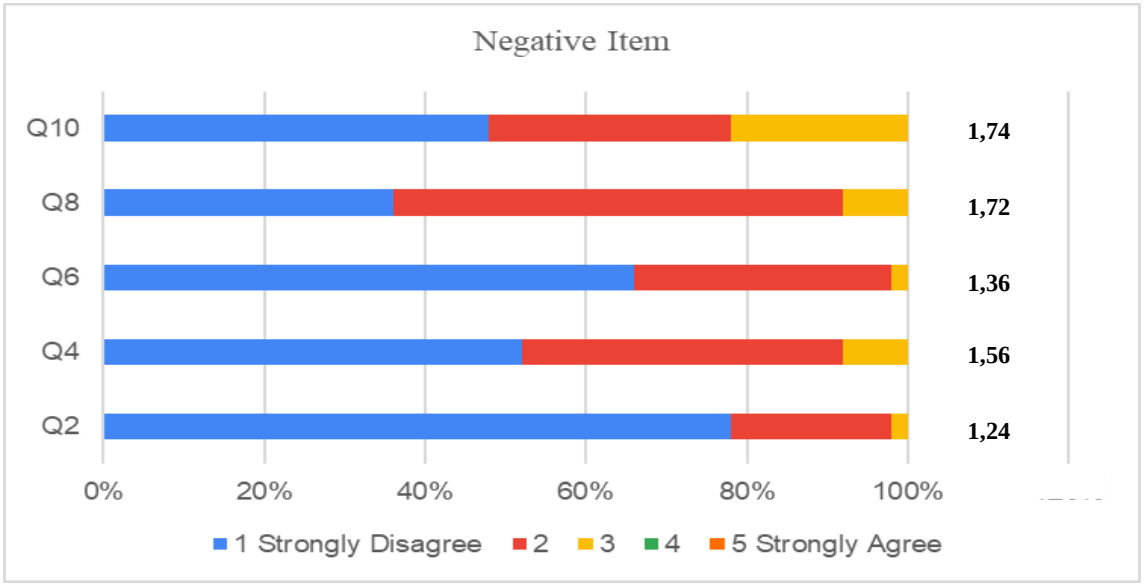

Figure 4. Percentage Spread and Average Value of Positive Negative Items

The SUS ratings may also be used to predict if a user would become a Promoter, Detractor, or Passive (Sauro & Lewis, 2011). According to Korneta (2018), the Net Promoter Scale (NPS) is determined based on the proportion of respondents who rate a product positively. If the evaluation score is less than 60, the category is Detractor; if the assessment score is between 90 and 100, the category is Promoter; and if the assessment score is between 70 and 80, the category is Passives. According to these five characteristics, promoters are happy with the system and will continue to use it. Meanwhile, a detractor is a user who has less experience with the given system or product and is more inclined to share unfavourable news

about the system if given the chance. As a result, it may be determined that UNJ LMS is included in the "Promoter" category, and users have already benefited from digital literacy.

CONCLUSIONS

The UNJ LMS's usability level was evaluated by users using the System Usability Scale (SUS) method, and the results were 83,25. The SUS score indicates that the system is in an "Acceptable" range based on the acceptability ranges. In terms of adjective ratings, the UNJ LMS falls within the "Excellent" category. Furthermore, according to the curved grade scale (CGS) standards, the UNJ LMS receives an "A" grade. Currently, users from the English Education Study Program's undergraduate and graduate programs rate the UNJ LMS positively. However, various upgrades should be implemented for users. These upgrades require UNJ LMS developers and management to regularly enhance the system in order to satisfy the demands and assist users. Some feature enhancements, such as virtual conferences and multimedia plugins for viewing and playing video, are required. Despite the score of the Net Promoter Scale being categorized as "Promoter" further research is necessarily conducted through the investigation of variables that affect the assessment based on expectations of different users (customer satisfaction). Since there must be different expectations from other departments within UNJ, it was also used as a basis to support decisions to improve the quality of service of each service system provided by the organization.

REFERENCES

Ali, W. (2020). Online and Remote Learning in Higher Education Institutes: A Necessity in light of COVID-19 Pandemic. *Higher Education Studies*, *10*(3), 16. https://doi.org/10.5539/hes.v10n3p16

Bangor, A., Staff, T., Kortum, P., Miller, J., & Staff, T. (2009). *Determining What Individual SUS Scores Mean : Adding an Adjective Rating Scale*. *4*(3), 114–123.

Borsci, S., Federici, S., Bacci, S., Gnaldi, M., & Bartolucci, F. (2015). *Assessing User Satisfaction in the Era of User Experience : Comparison of the SUS , UMUX and UMUX - LITE as a Function of Product Experience Assessing User Satisfaction in the Era of User Experience : Comparison of the SUS , UMUX , and UMUX-LITE as a Function of Product Experience*. *November*. https://doi.org/10.1080/10447318.2015.1064648

Brooke, J. (2013). *SUS : A Retrospective*. *January 2013*.

Coman, C., Țîru, L. G., Meseșan-Schmitz, L., Stanciu, C., & Bularca, M. C. (2020). Online teaching and learning in higher education during the coronavirus pandemic: Students' perspective. *Sustainability (Switzerland)*, *12*(24), 1–22. https://doi.org/10.3390/su122410367

Ependi, U., & Panjaitan, F. (2018). *System Usability Scale Antarmuka Palembang Guide Sebagai Media Pendukung Asian Games XVIII*. *3*(2).

Fearnley, M. R., & Amora, J. (2020). Learning Management System Adoption in Higher

Education Using the Extended Technology Acceptance Model Volume 8 – Issue 2 IAFOR Journal of Education: Technology in Education Volume 8 – Issue 2 IAFOR Journal of Education: Technology in Education Volume. *IAFOR Journal of Education: Technology in Education*, *8*(2), 89–106.

Grier, R. A., Bangor, A., Kortum, P., & Peres, S. C. (2016). *The System Usability Scale : Beyond Standard Usability Testing*. 187–191.

Korneta, P. (2018). *Net promoter score , growth , and profitability of transportation companies*. *54*(2), 136–148.

Leavitt, M. O. (2006). *Research-based web design & usability guidelines* (Vol. 2009, Issue July 12). http://www.usability.gov/pdfs/guidelines.html

Lewis, J. R. (2018). Measuring Perceived Usability : The CSUQ , SUS , and UMUX Measuring Perceived Usability : The CSUQ , SUS , and UMUX. *International Journal of Human–Computer Interaction*, *00*(00), 1–9. https://doi.org/10.1080/10447318.2017.1418805

N, I. A. H., Santoso, P. I., & Ferdiana, R. (2015). *Pengujian Usability Website Menggunakan System Usability Scale Website Usability Testing using System Usability Scale*. *17*(1), 31–38.

Safitri, S. T., Kusumawardani, D. M., Wiguna, C., Supriyadi, D., & Yulita, I. (2020). *MEASUREMENT OF VALIDITY AND RELIABILITY OF CUSTOMER SATISFACTION QUESTIONER in E-BOARDING APPICATIONS*. *16*(1), 1–6. https://doi.org/10.33480/pilar.v16i1.1069

Samir Abou El-Seoud, M., Taj-Eddin, I. A. T. F., Seddiek, N., El-Khouly, M. M., & Nosseir, A. (2014). E-learning and students' motivation: A research study on the effect of e-learning on higher education. *International Journal of Emerging Technologies in Learning*, *9*(4), 20–26. https://doi.org/10.3991/ijet.v9i4.3465

Sauro, J., & Lewis, J. R. (2011). *When Designing Usability Questionnaires , Does It Hurt to Be Positive ?* 2215–2223.

Sobaih, A. E. E., Hasanein, A. M., & Elnasr, A. E. A. (2020). Responses to COVID-19 in higher education: Social media usage for sustaining formal academic communication in developing countries. *Sustainability (Switzerland)*, *12*(16), 1–18. https://doi.org/10.3390/su12166520

Sugiyono. (2008). *Metode penelitian pendidikan: (pendekatan kuantitatif, kualitatif dan R \& D)*. Alfabeta. https://books.google.co.id/books?id=0xmCnQAACAAJ

Supriyadi, D., Thya Safitri, S., & Kristiyanto, D. Y. (2020). Higher Education e-Learning Usability Analysis Using System Usability Scale. *International Journal of Information System & Technology Akreditasi*, *4*(1), 436–446.