

# Comparison Between Usability and User Acceptance Testing on Educational Game Assessment

Nellya Anggun Vanesha<sup>[1]\*</sup>, Rizky<sup>[2]</sup>, Agus Purwanto<sup>[3]</sup>

Faculty of Computer Science<sup>[1], [2], [3]</sup>

Universitas Amikom Yogyakarta

Yogyakarta, Indonesia

Nellyaav12@students.amikom.ac.id<sup>[1]</sup>, samrizky@amikom.ac.id<sup>[2]</sup>, agus@amikom.ac.id<sup>[3]</sup>

**Abstract**— User Acceptance Testing (UAT) and Usability Testing are two methods commonly used in evaluating software or systems. UAT is concerned with overall system acceptance, while Usability Testing is specifically aimed at assessing the user's experience in interacting with the product. These two testing methods play an important role in ensuring the quality and user satisfaction of software and systems. Including being used to evaluate the Little Panda's Forest Animals game against 106 respondents consisting of two different campuses. The purpose of this research is to see the comparison between Usability Testing and User Acceptance Testing. With the research stages of literature review, questionnaire creation, data collection, data processing, and conclusions. The results of data processing show that there are differences in results where Usability Testing gets a lower score than User Acceptance Testing. Usability Testing results received an assessment range of 65 - 84 with the Usability statement being acceptable to users, while User Acceptance Testing received a range of 81% - 100% with the score interpretation criteria being very good.

**Kata Kunci**— Usability, SUS, UAT, Game

## I. INTRODUCTION

A game can be played with specific rules so that someone wins and loses, usually in a non-serious context or with the aim of refreshing. The gaming industry is a sector of the economy involved in developing, marketing, and selling video games. Currently, digital games have become part of a person's life activities. Mobile and computer technology support has made digital games a learning tool [1]. Elementary school children can digest material in educational games more efficiently than traditional textbooks [2]. One example of a game used for learning is Little Panda's Forest Animals.

Educational games are digital media that channel information through games. The quality of complex educational games, namely educational games, can be helpful as a learning medium for children to convey information, knowledge, or subject matter interactively and excitingly. An example of educational game testing is in a journal entitled Dancer vs Nondancer: Correlation of Controllers with Game Experience in Rhythm Games [3]. In evaluating educational games, the success of educational games is followed by disclosing most of the world's information and can be accessed directly [4]. One method used to measure game quality is the Webqual method, which measures website quality from three main dimensions: Usability Quality, Information Quality and Service Interaction

Quality [5]. Usability is the practice of testing applications or websites designed for users to see whether users can use them easily and have a good experience when using the application or website. Meanwhile, UAT (User Acceptance Testing) is the final test of the development of a product to validate that the system built meets user needs. However, the implementation still needs to be more optimal, so an evaluation is required to identify the UAT process.

Usability comes from the word usable, which means it can be used well. In general, usability optimizes interaction between users and the system, which can be done interactively. Usability is an essential aspect of the success of a website [6]. Usability is used to evaluate users' ease of use of a product [7]. Usability testing methods are divided into three, namely moderated and unmoderated, remote and in-person (face to face), and exploitive, assessment, and comparative.

Regarding usability, testing needs to be carried out, one of which is using the System Usability Scale (SUS) questionnaire. SUS is one of the most popular usability testing tools. SUS was developed by John Brooke in 1986. This SUS is a usability scale that is reliable, popular, effective, and cheap. SUS has several objectives, such as providing a measure of the user's subjective perception of the usability of a system and making it possible to carry out evaluations in a very short time. SUS is often used because it is effective and cheap, but the results are pretty good. Tests using the SUS method often use the questionnaire method.

User Acceptance Testing (UAT) is the final testing stage in software development, where users are given the opportunity to interact directly with the software before it is launched, ensuring that the system meets the user's needs and runs well. The definition of User Acceptance Testing (UAT) describes that User Acceptance Testing (UAT) is a test of a system carried out by users accepting application development and assuming that the user's needs have been met by the test results [8]. Meanwhile, User Acceptance Testing (UAT) is a test carried out by the final tester, where the user can be staff or company employees who interact directly with the system and can verify whether the existing functions are running. According to what is desired [9]. UAT aims to ensure that the tasks that have been developed align with user needs and expectations. The results of UAT are recorded in documents as evidence that users and their needs have accepted the application development are considered met based on test results [10]. This testing is

expected to test the game's internal and external usability and ensure that the game meets the criteria for education [11].

Research using the System Usability Scale was carried out by Bastiar and Gustalika in 2022 on the Design of Mobile Apps for Gamification of Folk Dance in Indonesia. This research used respondents with a total of 10 active students in Indonesia. The results show that the prototype evaluation using the System Usability Scale method shows that students agree with the existence of the folk dance gamification mobile application in Indonesia with an average result of 86.25% with an "Excellent" rating and a Letter Grade "A" [12]. Meanwhile, User Acceptance Testing was used in research entitled Implementation and User Acceptance Test (UAT) of the Integrated Library System (INLIS Lite) Application at Mts Negeri 7 Kuningan by Sambas and Ipan in 2022. The results of this research, namely the implementation of the INLIS Lite application, show that the application built meets the functional requirements for ease of use and usefulness. Using questionnaire testing which was taken from respondents totaling 308 students and 2 librarians. Where calculations in terms of design resulted in 76.2% of respondents answering agreeing with the appearance of the INLIS Lite application, then in terms of convenience 77.9% of respondents answered agreeing with using the INLIS Lite application. The final result of this research can be seen that the response from librarians and students to this application system based on the level of acceptance is strong, namely with a percentage of 77.2% [9].

In the field of product design, usability and user acceptance are two closely connected but different ideas, specifically when it relates to software or technology development. Users' willingness to use and accept a product is taken into consideration while determining its usability, while user acceptance focuses on how well, quickly, and satisfactorily users can do tasks using a product. It is crucial to compare usability to a user acceptability model for three reasons: Initially, Finding Gaps: It is possible to find any gaps between users' opinions of usability and their readiness to use a product by comparing usability to a user acceptance model. For instance, a product may still be rejected even when its efficiency and effectiveness are quite high if people find it difficult to use or if it doesn't meet their needs and expectations. Second, Holistic Evaluation: Someone may assess the product's overall effectiveness and potential for commercial success more comprehensively by attempting to take user acceptability and usability into account together. This can assist in guiding decisions about budget allocation, marketing tactics, and product design.

Third, Space for Improvement: Knowing more about how usability aspects affect user acceptance can highlight out areas in need of development. For example, if people find the product difficult to use even though it works well, this could mean that changes to the instructional materials or user interface design are necessary to improve user acceptance and enjoyment. All in all, a deeper comprehension of how designs affect users' perceptions and behaviors may be achieved by comparing usability to a user acceptance model. This, in addition, can lead to the development of more successful and user-friendly products. In this research, a comparison was carried out

between Usability and User Acceptance Testing in the educational game Little Panda's Forest Animals. The aim is to determine whether the game design is following user needs and the user experience of the game design. Then, the data is calculated using the SUS (System Usability Scale) method, which is under the application of the system usability scale method criteria [13]. The User Acceptance Testing method determines whether the system will be accepted, thereby providing customers confidence that the application meets all predetermined requirements and behaves as expected [14].

## II. RESEARCH METHODOLOGY

### A. Research stage.

This research went through several stages, as shown in Figure 1.

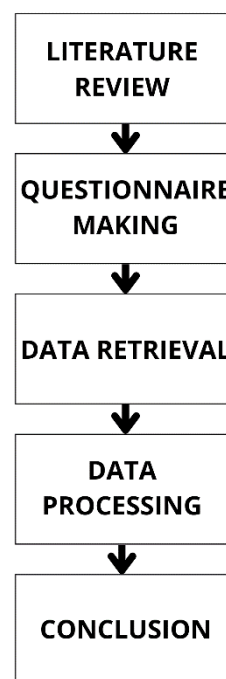


Figure 1. Research Stage

In Figure 1, there are five research stages. Starting from literature review, creating questionnaires, data collection, data processing, and conclusions. At the literature review stage, researchers collect references that will be used as references in compiling the journal. The second stage is that the researcher creates a questionnaire. The questionnaire aims to determine questions based on the SUS and UAT methods used to collect sample data. The third stage is the data collection process using a previously created questionnaire. The next stage is processing the data taken from the third stage. In this stage, the data collection results are processed using two testing methods. The first method uses SUS, and the second uses the UAT method. The final stage is the conclusion from the data processed in the fourth stage.

### B. Population dan Sampel

In evaluating the game Little Panda's Forest Animals, usability and user acceptance testing are required from a

population. A population is a group of people, animals, plants, or objects with specific characteristics that will be studied to be used as a basis for conclusions [15]. Samples from the population represent users who will later become respondents [16].

For this reason, the total N used in this research is 106, which will later be used as Usability and User Acceptance Testing respondents. After getting a sample to evaluate, the next step is determining the respondent category. The categories taken are based on the overall sample, campus (UNIPMA and AMIKOM), gender (Male and Female), and gaming experience (gamer and non-gamer). The following results were obtained.

1. Total (N) : 106 respondents.
2. Campus
  - a. UNIPMA : 54 respondents
  - b. AMIKOM : 52 respondents
3. Gender
  - a. Man : 44 respondents
  - b. Woman : 62 respondents
4. Game experience.
  - a. Gamer : 45 respondents
  - b. Non-gamer : 61 respondents

C. Data Collection Techniques

Data was collected using a questionnaire to collect numerical data, analyzing numerical data using SUS and UAT methods and then reporting the results of the analysis. Before collecting data through a questionnaire, respondents played the Little Panda's Forest Animals game for 5 to 10 minutes. Then, respondents were given time to interact with the game, followed by filling out a questionnaire by the respondent. This data was collected from September to October at two different campuses, namely UNIPMA and AMIKOM.

The questionnaire filled in is divided into two, namely as follows.

TABEL I. SYSTEM USABILITY SCALE (SUS) QUESTION LIST

No	Question	Score
1.	I'm thinking of using this game.	1-5
2.	This game is simple.	1-5
3.	I think this game is easy to use	1-5
4.	I need technical assistance in using this game	1-5
5.	I think that the functions in this game work well	1-5
6.	I feel there are several errors in this game	1-5
7.	I imagine that many of my friends will learn easily using this game	1-5
8.	I would argue that this application is impractical	1-5
9.	I feel very confident when playing this game	1-5
10.	I need to learn a lot before using this game	1-5

After respondents finished filling out the first questionnaire, which contained questions related to the System Usability Scale (SUS), they continued to fill out the second questionnaire, which included questions about UAT.

TABEL II. USER ACCEPTANCE TESTING QUESTION LIST

No	Question	Score
1	The appearance of this game is attractive	1-5
2	The menus in this game are easy to understand	1-5
3	The lesson material in this game is easy to understand	1-5
4	The case examples in this game help us understand the subject matter	1-5
5	The features in this game are structured	1-5
6	The evaluation section of the game helps gauge understanding of the material	1-5
7	This game can be used as a learning aid	1-5
8	This game runs well	1-5

Respondents fill out the questionnaire by filling in the score for each question using a five-point Likert-type scale [17][18]. Then, the results are collected in Microsoft Excel for data calculations.

TABEL III. LIKERT SCALE

Score	Information
1	Strongly disagree
2	Disagree
3	Neutral
4	Agree
5	Strongly agree

D. Data Processing and Analysis

After the questionnaire data given to respondents was collected, the data was divided according to the method. The calculation of the System Usability Scale (SUS) method is carried out using the conversion of respondent responses using [19]:

For odd statements, namely 1, 3, 5, 7, and 9, the score given by the respondent is reduced by 1, where Px is the odd question score.

$$Score\ SUS\ odd = \sum Px - 1 \quad (1)$$

For even statements, namely 2, 4, 6, 8, and 10, the scores given by respondents are used to subtract 5, where Pn is the even question score.

$$Score\ SUS\ even = \sum 5 - Pn \quad (2)$$

The conversion results are added up for each respondent and multiplied by 2.5 to get a value range between 0 – 100.

$$(\sum score\ odd + \sum score\ even) \times 2,5 \quad (3)$$

Once the score for each respondent is known, the next step is to

find the average score by adding up all the scores and dividing by the number of respondents. This calculation can be seen with the following formula:

$$\bar{X} = \frac{\sum x}{n} \quad (4)$$

Where  $\bar{X}$  is the average score,  $\sum x$  is the number of System Usability Scale scores, and  $n$  is the number of respondents. These results will obtain an average value from all respondents' score assessments. In this study, the following range of SUS assessment results was used [19]:

- Not acceptable = 00 – 64.
- Acceptable = 65 – 84.
- Perfect = 85 – 100.

Then, proceed with calculating User Acceptance Testing. The results of the UAT data obtained from filling out the questionnaire are multiplied by the weights in Table 3. Then, the average calculation is carried out by dividing the total weight of the respondents' scores by the total number of respondents [18]. So, you can find the percentage of each answer using the following formula [20]:

$$p = \frac{\text{Average}}{\text{Ideal Score}} \times 100\% \quad (5)$$

Where  $p$  is a percentage, the score is the total score obtained for each question, and the ideal score is the maximum. The calculation results will get a percentage value for each question. In this study, the following UAT score interpretation criteria were used [18]:

TABEL IV. UAT SCORE CRITERIA

Percentage	Result
0% - 20%	Very bad
21% - 40%	Bad
41% - 60%	Neutral
61% - 80%	Good
81% - 100 %	Very good

### III. RESULT AND ANALYSIS

#### A. System Usability Scale (SUS)

This research used data from 106 respondents consisting of users who had played the game Little Panda's Forest Animals. Respondents will answer ten questions in Table 1. The following are the data results from respondents:

TABEL V. SUS DATA

No	Question	Score				
		Strongly disagree	Disagree	Neutral	Agree	Strongly agree
1	I'm thinking of using this game.	5	9	42	32	18
2	This game is simple.	0	3	18	31	54
3	I think this game is easy to use	1	2	8	36	59

4	I need technical assistance in using this game	44	16	21	15	10
5	I think that the functions in this game work well	2	0	15	35	54
6	I feel there are several errors in this game	40	35	17	11	3
7	I imagine that many of my friends will learn easily using this game	6	5	26	40	29
8	I would argue that this application is impractical	54	36	10	4	2
9	I feel very confident when playing this game	1	4	15	34	52
10	I need to learn a lot before using this game	46	26	15	11	8

The results of the respondent's answers will be calculated using equations 1, 2, 3, and 4 for each category so that they will produce the following average score:

TABEL 6. SUS RESULT

No	Category	Average
1	Total	69,08
2	UNIPMA	67,31
3	AMIKOM	70,91
4	Man	69,14
5	Woman	69,03
6	Gamer	69,33
7	Non-Gamer	68,89

From the table above, it can be explained that the SUS calculation result for all respondents is 69.08. The results of the SUS calculation based on campus categories, namely UNIPMA with 54 respondents and AMIKOM with 52 respondents, are 67.31 and 70.91, respectively. The SUS value of 44 male respondents was 69.14, and the SUS value of 62 female respondents was 69.03. For the gamer category, 45 respondents obtained a result of 69.33 and 61 respondents got a result of 68.89 for non-gamers. The overall results and categories are in the assessment range of 65 – 84. So, it can be concluded that from the results of the SUS calculations, the game Little Panda's Forest Animals has been accepted by users.

#### B. USER ACCEPTANCE TESTING (UAT)

From 106 respondents consisting of users who have played the Little Panda's Forest Animals game, the 8 questions in Table 2 obtained the following results:

TABEL 7. UAT DATA

No	Question	Score				
		Strong disagree	Disagree	Neutral	Agree	Strongly agree
1	The appearance of this game is	2	1	14	32	57

	attractive					
2	The menus in this game are easy to understand	0	4	15	33	54
3	The lesson material in this game is easy to understand	0	1	15	34	56
4	The case examples in this game help us understand the subject matter	1	1	18	48	38
5	The features in this game are structured	1	1	27	41	36
6	The evaluation section of the game helps gauge understanding of the material	0	0	24	56	26
7	This game can be used as a learning aid	0	5	13	35	53
8	This game runs well	0	2	8	30	66

Table 7 shows the data results from filling out the UAT questionnaire. The UAT data results are multiplied by the weights in Table 3. Then, the percentage calculation is carried out using Equation 5, and the results for each category are obtained as follows:

TABEL 8. UAT RESULT

No	Question	Total	UNIPMA	AMIKOM	Man	Woman	Gamer	Non-gamer
1	The appearance of this game is attractive	87	86	87	82	89	85	87
2	The menus in this game are easy to understand	86	84	87	84	87	84	87
3	The lesson material in this game is easy to understand	87	87	88	86	88	88	87
4	The case examples in this game help us understand the subject matter	83	81	84	82	84	85	81
5	The features in this game are structured	81	82	80	77	84	80	81
6	The evaluation section of the game helps gauge understanding of the material	80	82	79	78	82	80	81
7	This game can be used as a learning aid	86	84	88	85	86	87	85
8	This game runs well	90	88	92	90	90	90	90

After calculating the percentage of User Acceptance Testing, the average results obtained for all questions from each category are as follows.

TABEL 9. UAT SCORE AVERAGE

No	Category	Average
1	Total	85
2	UNIPMA	84
3	AMIKOM	86
4	Man	83
5	Woman	86
6	Gamer	85
7	Non-Gamer	85

From the table above, it can be explained that the UAT calculation result for all respondents is 85%. The UAT calculation results based on the campus category, namely UNIPMA, which numbered 54 respondents, was 84%, while the AMIKOM campus category, which numbered 52 respondents, had a higher UAT result, namely 86%. The UAT score for male respondents, totaling 44 respondents, was 83%, and the UAT score for female respondents, totaling 62 respondents, was 86%. For the gamer category, 45 and 61 non-gamer respondents had the same UAT results, namely 85%. The overall results and categories are in the 80% -100 % assessment range. The game Little Panda's Forest Animals received an ideal interpretation score from the UAT calculations. The game Little Panda's Forest Animals has features following UAT's statement in its questionnaire.

IV. CONCLUSION

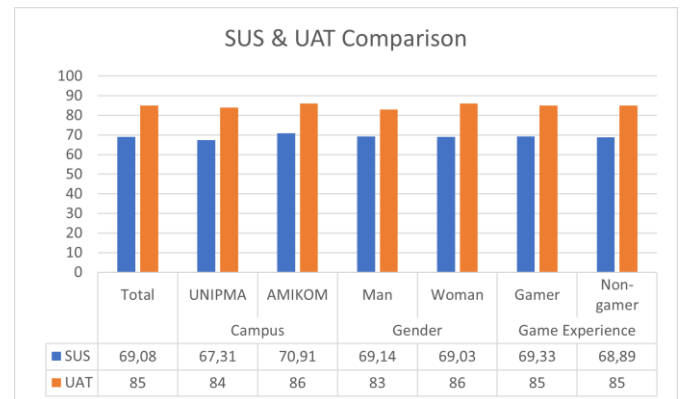


Figure 2. Comparison between SUS dan UAT

Based on research that has been carried out using the System Usability Scale and User Acceptance Testing methods, it can be seen in Figure 2 that UAT has higher results than Usability. The overall results and categories of the System Usability Scale method are in the assessment range of 65 – 84. So, it can be concluded that from the results of the SUS calculations, the game Little Panda's Forest Animals has acceptable Usability by users.

Meanwhile, using the User Acceptance Testing method, the percentage results for each category of all questions were 81% - 100%. In this range, the score interpretation criteria are excellent. From the results of the UAT calculations, the features in the Little Panda's Forest Animals game have been developed according to user needs and expectations and are running well.

There is a difference in assessment because the game being tested has several game mechanics in one game, so the usability

testing score gets a lower score than UAT. However, the games tested succeeded in meeting the basic needs of users in the learning process. Opportunities for further research are to try two different games and use varying measuring tools,

#### REFERENCES

- [1] A. Setiawan, H. Praherdhiono, and S. Suthoni, "Penggunaan Game Edukasi Digital Sebagai Sarana Pembelajaran Anak Usia Dini," *JINOTEP (Jurnal Inov. dan Teknol. Pembelajaran) Kaji. dan Ris. dalam Teknol. Pembelajaran*, vol. 6, no. 1, pp. 39–44, 2019, doi: 10.17977/um031v6i12019p039.
- [2] L. Lovita and T. H. Rochadiani, "Pengembangan Game Edukasi Visual Novel Mata Pelajaran Pendidikan Budi Pekerti Di SD Setia Bhakti," *Techno.Com*, vol. 22, no. 2, pp. 323–335, 2023, doi: 10.33633/tc.v22i2.7978.
- [3] M. Ni'maturahmah and . Rizky, "Dancer vs Non Dancer: Korelasi Kontroler Terhadap Game Experience Pada Rhythm Game," *Techno.Com*, vol. 21, no. 4, pp. 957–967, 2022, doi: 10.33633/tc.v21i4.6950.
- [4] J. F. Andry, K. Christianto, and F. R. Wilujeng, "Using Webqual 4.0 and Importance Performance Analysis to Evaluate E-Commerce Website," *J. Inf. Syst. Eng. Bus. Intell.*, vol. 5, no. 1, p. 23, 2019, doi: 10.20473/jisebi.5.1.23-31.
- [5] R. A. Nugraha, D. Andriyanto, D. Riana, and S. N. Khasanah, "Analysis of Factors Affecting Quality of corona.jatengprov.go.id Website towards User Satisfaction using Webqual 4.0 Method," *J. Phys. Conf. Ser.*, vol. 1641, no. 1, pp. 0–6, 2020, doi: 10.1088/1742-6596/1641/1/012066.
- [6] A. Sidik, "Penggunaan System Usability Scale (SUS) Sebagai Evaluasi Website Berita Mobile," *Technol. J. Ilm.*, vol. 9, no. 2, p. 83, 2018, doi: 10.31602/tji.v9i2.1371.
- [7] N. Huda, F. Habrizons, A. Satriawan, M. Iranda, and T. Pramuda, "Analisis Usability Testing Menggunakan Metode SUS (System Usability Scale) Terhadap Kepuasan Pengguna Aplikasi Shopee," *Simkom*, vol. 8, no. 2, pp. 208–220, 2023, doi: 10.51717/simkom.v8i2.158.
- [8] N. Aini, S. A. Wicaksono, and I. Arwani, "Pembangunan Sistem Informasi Perpustakaan Berbasis Web menggunakan Metode Rapid Application Development (RAD) (Studi pada : SMK Negeri 11 Malang)," *J. Pengemb. Teknol. Inf. dan Ilmu Komput.*, vol. 3, no. 9, pp. 8647–8655, 2019.
- [9] Sambas and Ipan Ripai, "IMPLEMENTASI DAN USER ACCEPTANCE TEST (UAT) APLIKASI INTEGRATED LIBRARY SYSTEM (INLIS Lite) DI MTs NEGERI 7 KUNINGAN," *ICT Learn.*, vol. 7, no. 1, 2022, doi: 10.33222/ictlearning.v6i1.2306.
- [10] V. M. Anjasmara and A. Hadi Sumitro, "Pengembangan Sistem Informasi Masjid Darul Arham Menggunakan Metode V-Model dan UAT (User Acceptance Testing)," *Inf. Syst. Educ. Prof. J. Inf. Syst.*, vol. 8, no. 1, p. 47, 2023, doi: 10.51211/isbi.v8i1.2443.
- [11] S. Syarif, T. Hasanuddin, and M. Hasnawi, "Perancangan Game Puzzle Labirin menggunakan Metode Game Development Life Cycle (GDLC) berbasis Unreal Engine," *Bul. Sist. Inf. dan Teknol. Islam*, vol. 3, no. 1, pp. 34–41, 2022, doi: 10.33096/busiti.v3i1.582.
- [12] A. R. Bahtiar and M. A. Gustalika, "Penerapan Metode System Usability Scale dalam Pengujian Rancangan Mobile Apps Gamification Tari Rakyat di Indonesia," *J. Media Inform. Budidarma*, vol. 6, no. 1, p. 491, 2022, doi: 10.30865/mib.v6i1.3510.
- [13] M. R. S. Sanjaya, A. Saputra, and D. Kurniawan, "Penerapan Metode System Usability Scale (Sus) Perangkat Lunak Daftar Hadir Di Pondok Pesantren Miftahul Jannah Berbasis Website," *J. Komput. Terap.*, vol. 7, no. 1, pp. 120–132, 2021, doi: 10.35143/jkt.v7i1.4578.
- [14] R. Suman and S. Sahibuddin, "User Acceptance Testing in Mobile Health Applications," pp. 145–149, 2019, doi: 10.1145/3322645.3322670.
- [15] J. Elektornik, S. Tinggi, and I. Ekonomi, "ejournal.stiegalileo.ac.id 30," pp. 30–38, 2018.
- [16] S. D. Purnamasari and F. Syakti, "Implementasi Usability Testing dalam Evaluasi Website Sekolah," *J. Sisfokom (Sistem Inf. dan Komputer)*, vol. 9, no. 3, pp. 420–426, 2020, doi: 10.32736/sisfokom.v9i3.1000.
- [17] A. Saputra, "Penerapan Usability pada Aplikasi PENTAS Dengan Menggunakan Metode System Usability Scale (SUS)," *JTIM J. Teknol. Inf. dan Multimed.*, vol. 1, no. 3, pp. 206–212, 2019, doi: 10.35746/jtim.v1i3.50.
- [18] J. Abraham and I. E. Ismail, "Unit Testing dan User Acceptance Testing pada Sistem Informasi Pelayan Kategorial Pelayanan Anak," *Repos. PNJ*, 2021.
- [19] A. W. Soejono, A. Setyanto, and A. F. Sofyan, "Evaluasi Usability Website UNRIYO Menggunakan System Usability Scale (Studi Kasus: Website UNRIYO)," *Respati*, vol. 13, no. 1, pp. 29–37, 2018, doi: 10.35842/jtir.v13i1.213.
- [20] I. Wahyudi and F. Alameka, "Analisis Blackbox Testing Dan User Acceptance Testing Terhadap Sistem Informasi Solusimedssosku," *J. Teknosains Kodepena*, vol. 04, no. 01, pp. 1–9, 2023.