Usability evaluation of a Gamification-based programming learning platform: Grasshopper

R Nuravianty¹, H B Santoso¹* and K Junus¹*

¹Faculty of Computer Science Universitas Indonesia Depok, Indonesia

* harrybs@cs.ui.ac.id, kasiyah@cs.ui.ac.id

Abstract. Online learning allows the learning process to be carried out anywhere and anytime. However, empirical studies report several obstacles that occur in the online learning process; one of which is the lack of student involvement. The learning materials of programming lessons generally have a low level of participation. Therefore, it requires a learning approach that is more attractive, easily understood by students, and promote engagement. Gamification-based programming learning platforms are widely available today. Research on determining the effectiveness of their use, investigating the system usability scale (SUS), perceived benefits, level of satisfaction, and user experience when using it is still limited. This study aims to evaluate a Grasshopper as one of the gamification-based programming learning platforms. Thirty-one respondents who had studied various programming languages at both the high school and university levels were involved in usability testing sessions. The results of the usability evaluation using SUS showed that the score was above average. The assessment of benefits and satisfaction also showed an average value of 8.6 which indicates that most respondents were satisfied with this application.

1. Introduction

During the COVID-19 pandemic, the use of e-learning has become a basic need in the learning process at all levels of education. However, several studies have revealed that there are various challenges or obstacles in conducting online learning, including the lack of student motivation in using online learning media [1][2], lack of student involvement [3][4], lack of face-to face interaction [5], and subject matter that is difficult to learn [1].

The lack of student involvement in the online learning process is caused by several factors, one of which is the lack of content that is attractive for students and the absence of rewards for students who contribute actively on the platform. Online educational sites such as codeacademy.com and khanacademy.org use game elements so that the users will be involved in the learning process. The use of game elements in a non-game context is called gamification [6]. Researchers have found that the use of gamification can increase student commitment in learning activities, students’ attendance and participation, students’ contribution in answering questions, and the percentage of students passing the course [7].

The difficulty level of a subject is one of the reasons students are less engaged in the learning process. Several studies on the effectiveness of the application of gamification were carried out in programming courses, because programming courses are subjects with a low level of student involvement and a high failure rate [1][8][9]. Research on the application of gamification elements has mostly carried out with the main focus on increasing the participation, involvement, and contribution of students. Gamification
design in a suitable learning system will improve student learning performance [10]. However, the application of gamification in online learning does not always result in students’ positive behavior, as found by research conducted by Kyewski and Krämer regarding the application of gamification in e-Learning courses. In fact, students felt that they were under external control and even social pressure[11]. In this study, Kyewski and Krämer indicated the need to evaluate the use of badges on the motivation, activities and performance of students in a more comprehensive way. This proves that the application of gamification design is not limited to operational requirements, but requires a deep understanding of human psychology [12].

Research on the application of gamification is mostly carried out using the experimental method by comparing one group of students that uses gamification and another group that does not use it. There are several researchers who have added questions as a form of evaluation of the results of the experiment. The theories used in research on the application of gamification elements are also diverse, such as the MDA Framework [1], flow theory [13], self-determination theory [14][15], self-determination theory and social comparison [11], goal setting theory [16], three dimension of engagement [17][18], and other theories that are in line with the research objectives. In addition to experiments and mixed methods, there are also researchers who only evaluate gamification on the Duolingo platform by using the Game Refinement Theory. The researchers measure Duolingo's refinement value by using data of users who have enrolled in Duolingo on the Duolingo website and the number of courses in each language [19].

Several studies related to the use of gamification in learning applications show that research on gamification-based platform usability analysis is rarely conducted in research related to gamification. Usability is a qualitative analysis that determines how easy it is for a user to use an application. Usability of user interface is part of the game aesthetics element [20], the other two are design and visibility. Aesthetics describe the user's expression, feelings, and emotional responses. Therefore, research related to usability analysis should be carried out in conjunction with an analysis of user benefits and satisfaction which can be extracted from experiments using a platform. The experience of respondents using a gamification-based platform can reveal the benefits obtained by users and the level of user satisfaction.

This study uses usability testing of the Grasshopper application to determine whether the application is feasible and has the potentials to help students learn programming languages. We used Grasshopper because it is free to access, has a very attractive design, and comes with brief and clear instructions. Good graphic design is important. Reading long sentences, paragraphs, and documents is difficult on the screen [21]. Thus, an inappropriate design will actually make users feel frustrated and lose motivation [22]. The method used in this research is the qualitative analysis method (observation) and quantitative analysis method (questionnaire). Observations were made of students' screen recordings when using the Grasshopper application. The questionnaire used consisted of three parts, namely, part one consisting of questions about the System Usability Scale (SUS) to measure the use of an application, part two consisting of questions about the benefits of the application felt by the respondents, and part three containing questions about user experience before and after using the application and the level of user satisfaction with the application.

The purpose of this research is to analyze the usability scale of the Grasshopper platform using the System Usability Scale (SUS), and to get the perceived benefits and satisfaction levels of the respondents involved by observation and analyzing the results of the questionnaires filled out by respondents after using Grasshopper.

2. Methods and Discussion
2.1. Respondents
Participants involved in this study were students who had been or were learning programming languages, both at the high school or vocational high school and university level. In the data collection process, thirty-one students met the criteria as shown in Table 1.
Table 1. Demographic information (n=31).

| Variable                          | Number | Variable                                      | Number |
|-----------------------------------|--------|-----------------------------------------------|--------|
| **Gender**                        |        | **Programming Language that is being or has been learned** |        |
| Female                            | 9      | HTML                                          | 23     |
| Male                              | 22     | CSS                                           | 16     |
|                                   |        | Java                                          | 21     |
| **Education Level**               |        | JavaScript                                     | 22     |
| High School                       | 19     | Python                                        | 11     |
| University                        | 12     | Visual Basic                                   | 6      |
|                                   |        | PHP                                           | 18     |
|                                   |        | Another one (ASP, Dev C and C++, Dart, C#, Angular, Node.js, Kotlin, Swift) | 10     |

2.2. Experimental Design

This study used the usability testing method with experiments carried out on the use of the Grasshopper application. All respondents used the same application, so there was no control group in this study. The assessment of these experiments was carried out qualitatively, both by observing and understanding the interpretation of the respondents' answers to the open-ended questions. In addition, there was a quantitative assessment of the value of System Usability Scale (SUS), the value of benefits, and the level of respondent satisfaction after using the application.

2.2.1. Observation. For the purposes of observation, scenarios were provided in the form of tasks that had to be completed by the respondents. In this section the screen recordings sent by respondents were saved, and the length of time it took respondents to complete each given task was recorded. The tasks given in each scenario were determined based on the usefulness of the Grasshopper application. The task scenarios given to the respondents can be seen in Table 2.

Table 2. Task Scenarios.

| Code  | Instructions                                                                 |
|-------|-----------------------------------------------------------------------------|
| TS1   | Log in to the Grasshopper application with your Google account, Facebook or Apple ID. |
| TS2   | Answer the question “Have you coded before?” When you first enter the Grasshopper application, then answer with “No, I'm new to coding” to enter the questions part in the “What's code?” |
| TS3   | Working on French Flag and Gabon Flag puzzles in the “Fundamentals” Menu. |
| TS4   | Take the “How Many Blue?” Quiz on the “Fundamental” Menu. |
| TS5   | Read and pay close attention to the material given on “Used a Function” in the “Fundamental” Menu. |
| TS6   | Open the Achievement page / menu to see “Concepts Unlocked”, “JavaScript Key Used”, “Day Coding Streak”! |
| TS7   | Create New Snippet on the “Code PlayGround” menu in the form of Indonesian Flag and Hello world! and Respondent’s name. |
| TS8   | Complete the “Drawing Boxes” and “Benin Flag” tasks in the “Practices” menu! |

Task scenarios in Table 2 were used for observational purposes. Observations were made to measure the level of efficiency in usability from the time required to work on a given task. Runtime assessed how
efficient a system was in user reusability. Measurement of the working time in the task scenario was carried out using the Windows media player, taking into account the start and end times. Time records from the use of the Grasshopper application done by the respondents can be seen in Table 3.

**Table 3. Time Records (in seconds).**

| Respondent | TS1 | TS2 | TS3 | TS4 | TS5 | TS6 | TS7 | TS8 |
|------------|-----|-----|-----|-----|-----|-----|-----|-----|
| R1         | -   | 32  | 81  | 6   | 4   | 19  | 95  | 40  |
| R2         | 31  | 56  | 99  | 8   | 12  | 18  | 114 | 48  |
| R3         | -   | -   | -   | -   | -   | 55  | 429 | 54  |
| R4         | 10  | 31  | 47  | 3   | 4   | 8   | 68  | 29  |
| R5         | 23  | 54  | 63  | 4   | 12  | 1   | 46  | 35  |
| R6         | -   | 45  | 67  | 5   | 3   | 6   | 120 | 52  |
| R7         | 25  | 70  | 125 | 10  | 12  | 19  | 444 | 36  |
| R8         | 18  | 36  | 70  | 22  | 219 | -   | 273 | 41  |
| R9         | -   | 33  | 58  | 3   | 10  | 8   | 248 | 29  |
| R10        | 45  | 46  | 73  | 9   | 28  | 13  | 430 | 39  |
| R11        | 8   | 43  | 57  | 4   | 10  | 25  | 113 | 48  |
| R12        | 10  | 48  | 48  | 10  | 3   | -   | -   | -   |
| R13        | 14  | 217 | 162 | 21  | 19  | 32  | -   | 111 |
| R14        | -   | 32  | 67  | 10  | 8   | 12  | 118 | 31  |
| R15        | 20  | 37  | 83  | 7   | 21  | 41  | 119 | -   |
| R16        | 17  | 80  | 81  | 7   | 12  | 16  | 208 | 37  |
| R17        | 2   | -   | 103 | 8   | 11  | 13  | 84  | 41  |
| R18        | 33  | 30  | 57  | 4   | 4   | 15  | 170 | 40  |
| R19        | 13  | 35  | 96  | 9   | 4   | 17  | 117 | 55  |
| R20        | -   | 10  | 76  | 7   | 3   | -   | 320 | 53  |
| R21        | 30  | 121 | 63  | 14  | 5   | 48  | 204 | 40  |
| R22        | 9   | 21  | 43  | 3   | 4   | 13  | 88  | 30  |
| R23        | 10  | 27  | 61  | 2   | 6   | 29  | 165 | 36  |
| R24        | 19  | 63  | 76  | 9   | 8   | 14  | 76  | 44  |
| R25        | 11  | 31  | 46  | 9   | 10  | 10  | 174 | 29  |
| R26        | -   | 26  | 54  | 3   | 3   | 6   | 80  | 36  |
| R27        | 19  | 27  | 71  | 4   | 2   | 5   | 102 | 41  |
| R28        | 12  | 46  | 65  | 10  | 4   | -   | 158 | 37  |
| R29        | 9   | 29  | 49  | 10  | 3   | -   | -   | -   |
| R30        | 9   | 30  | 54  | 8   | 40  | 12  | 151 | 37  |
| R31        | 17  | 67  | 152 | 43  | 8   | -   | 120 | 45  |
| Average    | 17,3| 49,1| 74,9| 9,1 | 16,4| 18,2| 172,6| 42,6|
Table 3 exhibited that on average all features could be done well by the respondents. Seven respondents recorded their activities in the TS2 section, so that the TS1 activities of these respondents could not be observed. Nevertheless, logging into the application was an activity student usually did.

2.2.2. Usability valuation of the Grasshopper Platform using SUS. In this section we used the SUS calculation adapted from the original by Sharfina and Santoso [23]. The list of questions in the research paper were translated into Indonesian and validated for use. The questionnaire used a Likert scale of 1-5, where 1: Strongly Disagree; 2: Disagree; 3: Neutral; 4: Agree; 5: Strongly Agree. The average SUS value obtained from the overall value given by the respondents is 69.27. A software product is considered to have a good usability value if the overall SUS value is equal to or above 68. In other words, the usability value of the Grasshopper application was quite above average.

2.2.3. Perception of the Benefits of the Grasshopper Platform. The questionnaire statements in this section were adapted from a research paper by Pinna et al. (2019) which has been confirmed by experts. The questionnaire uses a Likert scale of 1-5, where 1: Strongly Disagree; 2: Disagree; 3: Neutral; 4: Agree; 5: Strongly Agree. Table 4 shows the percentage of responses to questions about perceived benefits.

Table 4. Perceived Benefits.

| List of Questions                                      | 1 | 2 | 3   | 4   | 5   |
|-------------------------------------------------------|---|---|-----|-----|-----|
| Grasshopper helps you learn Programming subjects.     | - | - | 6,45% | 54,84% | 45,16% |
| Grasshopper helps you memorize programming concepts.  | - | - | 6,45% | 54,84% | 38,71% |
| Grasshopper can increase your competitiveness in      | - | - | 22,58% | 45,16% | 32,26% |
| learning Programming in the classroom.                |    |   |    |     |     |
| Using Grasshopper can reduce learning time on         | 3,22% | 9,68% | 32,26% | 41,93% | 12,91% |
| programming topics.                                   |    |   |    |     |     |
| Grasshopper has motivated you to study the            | 3,22% | - | 3,22% | 61,3% | 32,26% |
| programming topic.                                    |    |   |    |     |     |

2.2.4. Respondent Satisfaction Analysis of the Grasshopper Platform. This section contains a question "Rate your satisfaction level in using the Grasshopper application (from 1 to 10)"; and one closed question about whether the respondent would recommend the use of Grasshopper to others and why. The respondent assessment of their satisfaction level in using the Grasshopper application shows an average value of 8.6 which indicates that most respondents were satisfied with this application. 96.9% of the respondents would recommend using Grasshopper in learning programming languages. From a total of 30 respondents, one respondent did not recommend the use of Grasshopper because there was no Python programming language in Grasshopper.

3. Conclusion
To measure the perception of the benefits of gamification-based applications, the authors used a Likert scale 1-5 instrument to find out whether respondents agreed with the statements given in a questionnaire. From the data obtained it was found that most respondents agreed with the statements about the benefits of the platform they used. As for the level of satisfaction, 51.6% of the respondents gave a score of eight, 29% gave a score of nine, and 19.4% gave a score of ten on satisfaction in using the Grasshopper application, and 96.9% of the respondents would recommend the use of this application.

The analysis of the use of the Grasshopper application shows that it has a SUS score of 69.27. In other words, the value of the use of the Grasshopper application is above average. Thus, it can be concluded that the application can be accepted by users. Although most respondents benefited by using Grasshopper and were satisfied with it, they faced some obstacles, such language barrier of instruction used in the application. This is due to the fact that Grasshopper instruction language is in English only.
All the material in the Grasshopper application is in English. We recommend Grasshopper to add other languages of instruction to meet users’ need.

Acknowledgments

This research was supported by Hibah Publikasi Terindeks Internasional (PUTI) Prosiding 2020 at Universitas Indonesia (Number: NKB-846/UN2.RST/HKP.05.00/2020)

References

[1] B. Mar’uin, J. Frez, J. Cruz-Lemus, and M. Genero, “An Empirical Investigation on the Benefits of Gamification in Programming Courses,” *ACM Trans. Comput. Educ.*, vol. 19, no. 1, pp. 4:1-4:22, Nov. 2018.

[2] A. Bernik, G. Bubas, and D. Radosevic, “Measurement of the effects of e-learning courses gamification on motivation and satisfaction of students,” 2018 41st Int. Conv. Inf. Commun. Technol. Electron. Microelectron. MIPRO 2018 - Proc., pp. 806–811, 2018.

[3] R. G. de Fontes, D. S. Guerrero, and J. C. A. de Figueiredo, “Analyzing Gamification Impact on a Mastery Learning Introductory Programming Course,” in *Proceedings of the 50th ACM Technical Symposium on Computer Science Education*, 2019, pp. 400–406.

[4] G. Fraser, A. Gambi, M. Kreis, and J. M. Rojas, “Gamifying a Software Testing Course with Code Defenders,” in *Proceedings of the 50th ACM Technical Symposium on Computer Science Education*, 2019, pp. 571–577.

[5] S. Abu-Dawood, “The cognitive and social motivational affordances of gamification in E-Learning environment,” in *Proceedings - IEEE 16th International Conference on Advanced Learning Technologies, ICALT 2016*, pp. 373–375.

[6] S. Deterding, D. Dixon, R. Khaled, and L. Nacke, “From Game Design Elements to Gamefulness: Defining ‘Gamification,’” in *the 15th International Academic MindTrek Conference: Envisioning Future Media Environments*, 2011, pp. 9–15.

[7] D. Dicheva, C. Dichev, G. Agre, and G. Angelova, “Gamification in education: A systematic mapping study,” *Educ. Technol. Soc.*, vol. 18, no. 3, pp. 75–88, 2015.

[8] S. Azmi, N. A. Iahad, and N. Ahmad, “Attracting students’ engagement in programming courses with gamification,” in *2016 IEEE Conference on e-Learning, e-Management and e-Services, IC3e 2016*, 2017, pp. 112–115.

[9] A. M. Toda, R. M. C. do Carmo, A. P. da Silva, I. I. Bittencourt, and S. Isotani, “An approach for planning and deploying gamification concepts with social networks within educational contexts,” *Int. J. Inf. Manage.*, vol. 46, pp. 294–303, 2019.

[10] Y. Wang, P. Rajan, C. S. Sankar, and P. K. Raju, “Let Them Play: The Impact of Mechanics and Dynamics of a Serious Game on Student Perceptions of Learning Engagement,” *IEEE Trans. Learn. Technol.*, vol. 10, no. 4, pp. 514–525, 2017.

[11] E. Kyewski and N. C. Krämer, “To gamify or not to gamify? An experimental field study of the influence of badges on motivation, activity, and performance in an online learning course,” *Comput. Educ.*, vol. 118, no. November 2017, pp. 25–37, 2018.

[12] B. Morschheuser, L. Hassan, K. Werder, and J. Hamari, “How to design gamification? A method for engineering gamified software,” *Inf. Softw. Technol.*, vol. 95, pp. 219–237, 2018.

[13] T. Jagušt, I. Botički, and H.-J. So, “Examining competitive, collaborative and adaptive gamification in young learners’ math learning,” *Comput. Educ.*, vol. 125, pp. 444–457, 2018.

[14] G. Barata, S. Gama, J. Jorge, and D. Gonçalves, “Studying student differentiation in gamified education: A long-term study,” *Comput. Human Behav.*, vol. 71, pp. 550–585, 2017.

[15] Z. Zainuddin, “Students’ learning performance and perceived motivation in gamified flipped-class instruction,” *Comput. Educ.*, vol. 126, pp. 75–88, 2018.

[16] S. Nebel, S. Schneider, M. Beege, and G. D. Rey, “Leaderboards within educational videogames: The impact of difficulty, effort and gameplay,” *Comput. Educ.*, vol. 113, pp. 28–41, 2017.

[17] Ü. Çakıroğlu, B. Başbüyük, M. Güler, M. Atabay, and B. Yılmaz Memiş, “Gamifying an ICT
course: Influences on engagement and academic performance,” *Comput. Human Behav.*, vol. 69, pp. 98–107, 2017.

[18] L. Ding, C. M. Kim, and M. Orey, “Studies of student engagement in gamified online discussions,” *Comput. Educ.*, vol. 115, pp. 126–142, 2017.

[19] D. Huynh, L. Zuo, and H. Iida, “Analyzing gamification of ‘Duolingo’ with focus on its course structure,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 10056 LNCS, pp. 268–277, 2016.

[20] R. Hunicke, M. Leblanc, and R. Zubek, “MDA: A Formal Approach to Game Design and Game Research,” in *Proceedings of the AAAI Workshop on Challenges in Game Research*, 2004.

[21] J. Preece, Y. Rogers, and H. Sharp, *Interaction Design: Beyond Human-Computer Interaction*. John Wiley & Sons, 2006.

[22] B. J. Fogg, “A Behaviour Model for Persuasive Design,” in *Persuasive ’09: Proceedings of the 4th International Conference on Persuasive Technology*, 2009, pp. 1–7.

[23] Z. Sharfina and H. B. Santoso, “An Indonesian adaptation of the System Usability Scale (SUS),” *2016 Int. Conf. Adv. Comput. Sci. Inf. Syst. ICACISIS 2016*, pp. 145–148, 2017.