User interface evaluation of a real-time gimmick tracking

N R Nurwulan¹ and K F Zahirah¹

¹Department of Industrial Engineering, Sampoerna University
Jl. Raya Pasar Minggu Kav. 46, Jakarta, 12780, Indonesia
E-mail: nurul.nurwulan@sampoernauniversity.ac.id

Abstract. This paper presents the usability evaluation of a gimmick tracking application designed by students in the Human-Computer Interaction class to solve the problem in one of the biggest food companies in Indonesia. Gimmick tracking is important for the company to make sure the gimmick can be transferred to the customers with minimum-to-none problems. The gimmick tracking application should be user-friendly for the users to make sure the application meets the purpose of its development. Quantitative and qualitative usability evaluation using the system usability scale (SUS) questionnaire and think-aloud method were used to evaluate how effective is the real-time gimmick tracking application. A total of 12 subjects with an age range from 18 to 27 from the grocery stores in Malang were recruited to participate in this study. The SUS questionnaire was used to measure the hypotheses related to learnability, efficiency, memorability, errors, user satisfaction, and usability. The qualitative data such as recorded videos and notes were transcribed into the text for further analysis. The SUS score of the design of the gimmick tracking application was 70.83, this score is higher than 55 percent of other applications. Thus, it can be concluded that our design generally has fulfilled all usability requirements.

1. Introduction
PT XYZ is a multinational food and beverage company with main businesses in Indonesia are in waters and specialized nutrition products. As one of the biggest food and beverage companies in Indonesia, real-time gimmick tracking is important for PT XYZ Indonesia to ensure the gimmick could be delivered to the customers properly. Gimmick includes products and gifts for customers in terms of promoting the main products. Based on the information obtained from the interview with the company, tracking the gimmick is important for PT XYZ because it spends a 12% cost ratio that is worth 60 billion rupiahs per year for gimmick-related promotion. However, there is an average of 7% gap between the produced and delivered gimmick annually, this gap worth about 720 million rupiahs.

Indonesia is a growing market for mobile applications due to the number of the population. In addition, mobile phones cannot be separated from human daily lives [1]. A good gimmick tracking application should be able to connect the data of the retailers and PT XYZ in real-time to minimize the mismatch between produced and delivered gimmicks. Thus, the app needs to be user friendly for the users in the retail stores. A poorly designed user interface may create more problems in documenting the report. Functionality and good visuals are important in enforcing the objective of the application and easing the users as it facilitates interactions between the user and the application.

Usability is the most widely used to evaluate the user interface of the application [2]. It simulates the interaction between the user and the application under controlled conditions. The aim of this study was
to conduct a user interface evaluation using usability testing for the real-time gimmick tracking application of PT XYZ Indonesia that was designed by the students as a project for the Human-Computer Interaction class. This study is important to ensure the application could help the company in minimizing the mismatch of produced and delivered gimmick. As for the students, it is beneficial for them to know the proper methods in designing a mobile application that considers both user interface and user experience. Students were asked to solve the problems faced by PT XYZ Indonesia using their systems thinking skills. Systems thinking is essential in STEM education because it helps students to understand complex systems in order to solve problems [3, 4]. In systems thinking, students are fostered to see the big ideas of the problem, conceptualize, design, and implement the solutions to the problems [4].

The organization of this paper is prepared as follows. The Methodology section provides the experiment protocol and a brief description of the theoretical background behind this study. The results and discussion of this study are presented in section 3. The conclusion is presented in section 4.

2. Method

The first step in designing the gimmick tracking application was doing interviews and observations related to the situation in PT XYZ Indonesia. Then, students designed the application using a paper mockup in detail. After the design was completed, students did the usability evaluation of the gimmick tracking application in several retail stores that sell the products of PT XYZ Indonesia. There are main factors in usability evaluation, namely learnability, efficiency, memorability, errors, satisfaction, and usability [5, 6]. Thus, we developed the hypotheses of the study based on each factor as can be seen in Figure 1. The hypotheses are: H1) Users easily learn in operating the application; H2) Users quickly perform the given tasks; H3) Users able to retain their knowledge well; H4) Users make corrective errors; and H5) Users are satisfied with the application. Using the hypotheses above, we collected data from a total of 12 store attendants aged 18 to 27 who were recruited from the local grocery stores that spread around 5 districts in Malang, Indonesia. In usability evaluation, 12 subjects are enough to get all the needed data, since the minimum of subjects to get 90% of problem discovery is 7 subjects [7].

Figure 1. Hypotheses of the study
Figure 2. Main menu of the gimmick tracking application: (a) frontpage; (b) main menu; (c) report menu; (d) promotion menu; (e) notification menu; (f) profile menu
2.1. System usability scale
The system usability scale (SUS) questionnaire consists of a brief background survey, agreement to participate in the experiment, SUS survey, and objective rating scale. In this study, we modified the SUS questionnaire of Kortum and Bangor [7].

2.2. Think-aloud
The think-aloud (TA) protocol was chosen as the method in this study because we wanted to test how user-friendly is the application for the users. In this session, the experiment was recorded for further analysis. In this session, the subjects were asked to explain how to use the application loudly [8]. The experimenter reminded the subjects to keep talking whenever they stopped talking for more than 1 minute. The purpose of reminding the subjects was to let them keep on thinking actively in order to figure out how to use the application correctly. The interface of the gimmick tracking application that we designed can be seen in Figure 2. When the users open the application, the frontpage will show up to let the user either register or log in if they already have the account. After the user login, the main menu consists of report, promotion, notification, and profile will show up. The report menu will let the user register the gimmicks that have been received and delivered to the customers. This menu enables both companies and retailers to track and reduce the gimmicks lost in the system. The promotion menu shows the recent promotions of each product of PT XYZ Indonesia. The notification menu notifies the user when the gimmicks have been delivered to the store. Whereas the profile menu enables the user to update the information of the stores.

2.3. Data analysis
2.3.1. Quantitative Data
The questions in the SUS questionnaire were categorized into 10 main categories, namely: 1) I think that I would like to use this system frequently; 2) I found the system unnecessarily complex; 3) I thought the system was easy to use; 4) I think that I would need the support of a technical person to be able to use this system; 5) I found the various functions in this system were well integrated; 6) I thought there was too much inconsistency in this system; 7) I would imagine that most people would learn to use this system very quickly; 8) I found the system very awkward to use; 9) I felt very confident using the system; 10) I needed to learn a lot of things before I could get going with this system.

Each item’s score contribution was then determined in measuring the SUS score. The contribution of the score for positively worded items was the position of the scale minus 1. The contribution of the score for negative-worded items was 5 minus the position of the scale. Then, the sum of item score contributions was multiplied by 2.5 to get the overall SUS score, as can be seen in equation 1.

\[
SUS\ score = [(Q_1 - 1) + (5 - Q_2) + (Q_3 - 1) + (5 - Q_4) + \cdots + (Q_{n-1} - 1) + (5 - Q_n)] \times 2.5
\]  

where \(Q\) is the score of the \(n\)-th question.

Therefore, in a 2.5-point increment, SUS scores ranged from 0 to 100 [9]. The average SUS score is 68. Thus, the SUS score above 68 is considered above average and any score below 68 is below average. To interpret the SUS score, we converted the score to a percentile rank using the normalizing process. For example, a 55 SUS score was converted to a 20 percent percentile rank, which implies that the perceived usability of a SUS score of 55 is higher than 20 percent of other products. Whereas for the adjective question, each option was quantified in a 2.5-point increment with a score range from 0 to 10. The point for the adjective question was added to the user satisfaction aspect.
2.3.2. Qualitative Data
The recorded videos, notes, and other supporting documents were collected and organized for analysis. The recorded videos were transcribed into text, while notes and other observations were inserted into the transcription. The transcription was divided into segments that were then indexed by information (subject, problem, date) and by model categories. A coding scheme was defined based on the psychological model and verbalization theory. The coding scheme from the psychological model was done by taking every sub-process distinguished in the model and state how the processes were expected to appear. For example, the sub-process named ‘guessing’ was assigned when the subject mentioned: ‘Could it be X?’, ‘Maybe it is X’ or ‘Let us try X’.

The next step was comparing the protocols with procedural models. Each segment in each protocol should fit within the model. If a segment did not fit into the model, it was marked as a deviation. The differences between the coded protocols and the predictions from a process model include unpredicted processes, absence of predicted processes, and unpredicted sequences. Finally, the result of the protocol analysis can be reported. The documents to be documented at the completion of the analysis process are the detailed model and the coded protocols [10].

3. Results and Discussion
Based on experimental data we collected in several districts of Malang, Indonesia from March to April 2020, the results of the usability of the gimmick tracking application could be described as follows.

3.1. Quantitative data
The usability variables namely learnability, efficiency, memorability, user satisfaction, and usability were quantified based on the collected SUS questionnaires. The SPSS software was used to calculate the independent t-test, normality test, multicollinearity test, and Pearson’s correlation. Table 1 shows the descriptive statistics of each usability variable.

| Variable    | Min | Max  | Mean  | Standard Deviation |
|-------------|-----|------|-------|--------------------|
| Learnability| 10.00 | 27.50 | 19.79 | 4.70               |
| Efficiency  | 15.00 | 20.00 | 16.25 | 1.69               |
| Memorability| 2.50  | 10.00 | 7.29  | 2.49               |
| User satisfaction| 10.00 | 37.50 | 27.50 | 7.83               |
| Usability   | 40.00 | 90.00 | 70.83 | 13.03              |
| Error       | 0.00  | 2.00  | 0.50  | 0.67               |

The independent t-test was used to see whether the individual characteristics affect the results of the SUS questionnaire. Based on the calculation, our data are homogenous and thus, the individual characteristics do not affect the SUS score. As for the normality test, it was used to evaluate whether the data were collected from a normally distributed population. We used the Shapiro-Wilk test for the normality test and the result showed our data are indeed normal. To ensure the precision of the estimated variables, the variance inflation factor (VIF) of the multicollinearity test must be less than 10 and the value of tolerance must be higher than 0.01. In our data, the values of VIF are all lower than 10 and the values of the tolerance are all higher than 0.01. Therefore, there is no multicollinearity detected in our dataset. Pearson’s correlation test was used to see if there is any correlation between the usability variables. Based on our calculation, usability highly correlates (p < 0.05) with learnability, memorability, and user satisfaction with a correlation value of 0.88, 0.87, and 0.91, respectively.
3.2. Qualitative data
There are five main things that we found from the think-aloud procedure, namely: 1) Suggestive inputs; 2) Error and corrective actions; 3) Users did not go through the steps one-by-one; 4) Users did not think-aloud when reading the features; and 5) Users did not try anything in a section. The users gave several suggestive inputs such as adding a sorting option and promotion code in the notification menu, add a confirmation message for updating the profile, and add an option to show password. The errors done by the users in the usability experiment could also be used to improve the application by adding new options to prevent errors. As for the last three findings related to the think-aloud procedure, we had to always remind the users to do the experiment based on instructions.

3.3. Discussion
The Real-Time Gimmick Tracking application was designed based on the needs and flows of PT XYZ Indonesia’s routines in order to minimize the gap between produced and delivered gimmick. The students designed the gimmick tracking application and then performed the usability evaluation to see the relevance of the design to the real problem faced by PT XYZ Indonesia. The usability evaluation is important to ensure the gimmick tracking application can be used by the retailers of PT XYZ’s products. The analysis of the collected quantitative data showed that the application is above average and is higher than 55 percent of other products with a SUS score of 70.83. To make sure the collected quantitative data were valid for analysis, normality, independence, and correlation tests were done using the SPSS. The results from SPSS showed that the data were normal (p > 0.05 using the Shapiro-Wilk Test), independent (p > 0.01 for Tolerance, VIF < 10), and the correlation between the dependent variable (Usability) with the factors (Learnability, Memorability, and User Satisfaction) is high (Pearson’s correlation coefficient > 0.75). Therefore, the collected data in this study were valid. Moreover, the independent t-test showed that all factors (age, gender, education background) were homogenous (p > 0.05) and did not have significant differences between the two separated groups (p > 0.05 for Equality of Means). This means that the subjects’ characteristics did not affect the results of the usability evaluation. Based on the evaluation, it can be said that all 5 hypotheses were accepted, meaning users easily learn in operating the application, users quickly perform the given tasks, users able to retain their knowledge well, users make corrective errors, users are satisfied with the application. Nevertheless, suggestive inputs from the subjects obtained for the think-aloud session are also important for the improvement of the real-time gimmick tracking application.

The findings in this study are in agreement with previous studies [11-13]. The suggestive input taken from the Learnability aspect was to add labels to the icons to make learning easier [11] since one of the participants forgot the name of the pages and it slowed him down in using the application. For the Memorability aspect, the suggestive input was to give confirmation to the participant after any action to ensure the system constantly gives feedback on participants’ actions so they can always tell the state of the system [12]. For the Error aspect, there were several errors made by the participants during the testing. The system designer can either eliminate error-prone conditions or check and present participants with a confirmation option before they do the action [13]. In this study, we added confirmation before participants did the action.

The limitation of this study is that we only designed the gimmick tracking application and tested the design. However, we did not develop the mobile application. We only gave suggestions to PT XYZ Indonesia based on the findings of this study. It is because this study was done to support and foster students to be creative and innovative in solving the real-word case by using the knowledge they have received in the class. Using systems thinking, students can give valuable suggestions that potentially could solve the lost gimmicks in PT XYZ Indonesia.

This class project helped students in understanding the important knowledge of Human-Computer Interaction that could not be delivered in the lecture session in the class. Past studies reported that project-based learning could effectively motivate low performing students to try to understand the lectures better [14-16]. In project-based learning, students are trained to use critical thinking in analyzing, evaluating the relevance of the information gathered, and interpret them to solve problems.
Project-based learning is suitable for promoting a deeper level of understanding in learning since it enables peer-mentoring, leadership skills, and critical thinking that might not be achieved from the lectures by the teacher [15]. The role of the teacher in this project was guiding the students in designing the gimmick tracking application’s mockup and conducting the usability evaluation experiment. The systems thinking approach was chosen to enable students in focusing on the big picture of the problems [4]. Then, students were required to connect knowledge related to Human-Computer Interaction lectures to solve the problems. The focus on the big picture could help students solve the problems in a more effective and efficient manner.

4. Conclusion

The gimmick tracking application was made to reduce the gap between the produced and executed gimmick. This study evaluated the user interface of the real-time gimmick tracking application to get a better understanding of how users perceive the application. Students in the Human-Computer Interaction class made the design of the gimmick tracking application using paper mock-ups and then conducted the usability evaluation of the design to 12 store attendants in Malang, Indonesia. The evaluation analysis showed that the application is above average with a SUS score of 70.83. However, there were several suggestive inputs to improve the application given by the participants from the think-aloud session. The independent t-test from SPSS showed that all factors (Age, Gender, Highest Education) were homogenous and did not have significant differences between the two separated groups. The analysis from SPSS also showed that the data were normal, independent, and the correlation between the dependent variable with the factors was high. Thus, it can be concluded that all hypotheses were accepted. In this study, we did not develop the gimmick tracking application. We only designed the application and then gave suggestions to the company based on the results of the usability evaluation. Future works might consider developing a real mobile application to evaluate the usability of the application. In conclusion, the class project was able to help students in understanding the course materials better by applying their knowledge in the lecture sessions to solve real-world problems.

References

[1] Nurwulan N R, Iridiastadi H and Jiang B C 2015 A review of the effect on postural stability while using mobile phone The 4th Int. Conf. on Healthcare Ergonomics and Patient Safety (HEPS) Taipei, Taiwan 101-108
[2] Lewis J R, Utesch B S and Maher D E 2015 Measuring perceived usability: The SUS, UMUX-LITE, and Alt Usability International Journal of Human-Computer Interaction 31(8) 496-505
[3] York S, Lavi R, Dori Y J and Orgill M 2019 Application of systems thinking in STEM education J. Chem. Educ 96(12) 2742-2751
[4] Chalmers C and Nason R 2017 Systems Thinking Approach to Robotics Curriculum in Schools In: Khine, M S (ed.) Robotics in STEM education: Redesigning the learning experience (Switzerland: Springer) p 33-57
[5] Creswell J W 2018 Research Design: Qualitative, Quantitative, and Mixed Methods Approaches 5th ed (California: Sage Publications, Inc)
[6] Khajouei R, Esfahani M Z and Jahani Y 2016 Comparison of heuristic and cognitive walkthrough usability evaluation methods for evaluating health information systems J. Am. Med. Inform. Assoc 24(1) 55-60
[7] Kortum P and Bangor A 2013 Usability ratings for everyday products measured with the system usability scale International Journal of Human-Computer Interaction 29(2) 67-76
[8] Eccles D W and Arsal G 2017 The think aloud method: what is it and how do I use it? Qual. Res. Sport Exerc. Health 9(4) 514-531
[9] Lewis J R and Sauro J 2017 Can I leave this one out? The effect of dropping an item for the SUS Journal of Usability Studies 13(1) 38-46
[10] Bai B 2018 Understanding primary school students’ use of self-regulated writing strategies through think-aloud protocols System 78 15-26
[11] Inal Y 2018 University students’ heuristic usability inspection of the national library of Turkey website Aslib J. Inf. Manag. 70(1) 66-77

[12] Hussain A, Mkpojiogu E O C and Fazillah M K 2016 A systematic review on usability evaluation methods for m-commerce apps J. Telecommun. Electron. Comput. Eng. 8(10) 29-34

[13] Nielsen J and Loranger H 2006 Prioritizing Web Usability (California: New Riders Press)

[14] Mutakinati L, Anwari I and Yoshisuke K 2018 Analysis of students’ critical thinking skill of middle school through STEM education project-based learning Jurnal Pendidikan IPA Indonesia 7(1) 54-65

[15] Hall A and Miro D 2016 A study of student engagement in project-based learning across multiple approaches to STEM education programs Sch. Sci. Math. 116(6) 310-319

[16] Kellgren A C, Parker C E, Blustein D L and Barnett M 2016 Innovations and challenges in project-based STEM education: Lessons from ITEST Journal of Science Education and Technology 25 825-832