

Mobile Flight and Hotel Booking Application: A Heuristic and UX Test

Azham Hussain, Rahma Fitria

Human-Centered Computing Research Lab, School of Computing, Universiti Utara Malaysia, Malaysia.
azham.h@uum.edu.my

Abstract—The purpose of this study is to present the usability evaluation of the mobile flight and hotel booking application. Currently, mobile applications have offered the easiness to users in order to book flight and hotel. Unfortunately, there is a lack of study on usability that focuses on mobile booking application, while many people start to do all the transactions on their mobile phone. This study will conduct two combined usability evaluation methods which are heuristic and UX test. Ticket.com will be evaluated as it is a critical application that can represent the study of this usability testing. Those selected usability evaluation methods are to evaluate mobile application based on usability experts and end users to get feedbacks and propose solutions and recommendations to improve the application. Observation and questionnaire will be used to measure the metric of effectiveness, efficiency and satisfaction. The heuristic is conducted by three usability experts. Whereas, the UX test participants are ten persons that will be randomly selected either from novice users or the one who experience purchasing flight ticket and hotel reservation on mobile. The result of the study revealed that the application is not easy-to-use and inconsistent as it has been proven by usability score of this application is 66% or below average. It is expected that this study can be adopted by developers and usability practitioners to deliver a user-friendly mobile application that leads to high customer satisfaction.

Index Terms—Flight and Hotel, Heuristic Evaluation, Mobile Booking, Usability Testing, UX Test.

I. INTRODUCTION

The Internet usage has entirely revolutionised the behaviour of people in the way of purchasing goods and services. Nowadays, people begin to shop online as their life become very busy to be able to shop in-store. Hence, the development of mobile phone has rapidly grown to satisfy the need of human being as they start using mobile phone habitually in their daily life. Mobile phone emerges as one of the devices that people always use in their daily life. Purchasing flight ticket and hotel booking on the mobile phone has become trends among people who attach very much on the easiness of purchasing that produce a paperless ticket. According to Nielsen global e-commerce report in 2017, more than a half of global online purchasing on fashion products accounted for 58%, travel products or services represented an average 55%, Book, music & stationary represented 50%, IT and mobile accounted for 43% and event tickets is 41% of the total global respondents [1].

The growth of current technology on the mobile phone gives a big opportunity for airlines, travel and tourism companies to attract customers by offering the easiness of purchasing on the mobile application [2]. According to Nielsen Mobile Wallet Syndicated Report in 2016, the vast majority (76%) of Canada smartphone owners have used their

mobile phone in purchasing-related activity [3]. A Bronto report in 2016 also highlights that 64% of Americans are shopping more often on their mobile phones [4]. From the combination surveys above shows that purchasing travel products or service is the second most likely purchased product/service in online shopping that attract smartphone users. The tremendous accomplishment of these companies can be interpreted by developing their mobile application that is user-friendly. A lot of similar application has been introduced in the market in order to book flight ticket or hotel. However, that application that has poor usability will not attract the users or customers to use those applications in the future.

Many kinds of research carried out usability study on mobile applications, but a few research that focuses on mobile booking applications [5], [6], [7]–[9]. There is two-closely-related study that focuses on mobile flight booking applications [10], [11]. However, none all of them combined two usability evaluation methods (UEMs) in evaluating mobile booking application. The previous study on mobile flight booking application applied inquiry methods which are the questionnaire and unstructured interview. In 2005, Holzinger conducted a study which showed that questionnaire is conducted only to collect the opinions of the users about the interface. In addition, It has argued the users' answers on the questionnaire in which it really reflects where the users answer not really follow their true feeling [12]. Another problem, when the user said they are satisfied and content about the application, does it mean they were able to efficiently complete the task given, or that though the application failed them, and are these measures really quantifiable?. Meanwhile, the unstructured interview will cause the difficulty to interpret result [12].

In contrast, the proposed study is to conduct two usability evaluation methods (UEMs) which are heuristic and UX test in order to find and analyse issues based on usability experts and end users. The metrics used in this UX test are to measure how effective, efficient and satisfy the application is [13]. Furthermore, the issues found in the mobile application can be identified and analysed in order to propose the solutions and recommendations on the application. As the result, all the issues and feedbacks from usability experts and users will elevate the usability of this mobile flight and hotel booking application.

II. RELATED WORK

Usability can be defined as the extent to which a system or product is used by specified users in order to achieve the effectiveness, efficiency and user satisfaction [14]. Usability evaluation is an important parameter in software

development process in order to produce a usable product [15]. Usability evaluation methods (UEMs) is divided into inquiry methods, inspection methods and formal user testing [16]. Inquiry methods can be included in focus groups, interviews, questionnaires and surveys. For inspection methods are heuristic evaluation (HE) and cognitive walkthrough (CW). Finally, methods of formal user testing involve employing user experiments.

Heuristic evaluation is one of the inspection techniques that is conducted by usability expert in order to find some issues on the application by following a set of usability guidelines [15, N and M]. This usability evaluation method (UEM) propose a good way to decrease cost and produce the results in short time where there is no time-consuming in recruiting user [15]. However, the results of the heuristic evaluation cannot be completely trusted where there are no actual users involved [15].

The previous study has been conducted the heuristic evaluation on PDA-based supermarket application. The authors have applied 8 mobile heuristics and 10 Nielsen's heuristics [18]. Another previous study on the implementation of the heuristic evaluation conducted on mobile learning course content application (MOSAD) [19]. In contrast, this study will conduct heuristics evaluation by following 12 touchscreen mobile heuristics [20]. 10 Nielsen's heuristics is also to be used for references to get more details explanation about heuristics evaluation principles. The 12 touchscreen mobile heuristics is the update heuristics principles of 10 Nielsen's heuristics that focus only on the touchscreen-based mobile device. Furthermore, The issues found in the heuristic evaluation will be used to design test scenario.

User experience test (UX) can be defined as a users' perceptions and responses after using the product or system [21]. UX test allows multiple observers to be involved in evaluation at one session. It is also conducted in a controlled environment [12]. The previous study on UX test has been implemented on the mobile wire by comparing laboratory and field testing test [22]. The testing is conducted in the laboratory environment and in the field where the user can sit or stand during the test. However, this study will conduct UX test in two locations and various users background. As there is no usability laboratory in UUM campus, the test will be conducted in a room or laboratory-like area. The number of users is around 5 to 10 persons [22]. The user will be given several tasks to be completed and at the same time, the usability specialist will do the observation to measure effectiveness and efficiency metric. In the end of the test, the user will be given post-test questionnaire in order to measure satisfaction metric.

The previous work that has been conducted on touchscreen mobile flight booking application is applying inquiry method as the chosen usability evaluation methods (UEMs) used [11], [23]. The authors conducted inquiry method by performing unstructured interviews with 20 interviewees and then distributed questionnaires to them.

Another closely-related study is mobile air ticket booking on the classic mobile phone with tiny screen size [10]. The author proposed a designed prototype of mobile air ticket booking is tested by applying UX test. Unfortunately, the design prototype was not able to be tested to real users as during the time published the design prototype not completely done.

Table 1
The previous study on mobile flight booking application

No	Usability Evaluation Methods (UEMs)	Author(s)
1	Inquiry Method (Unstructured Interview & Questionnaire)	[11], [23]
2	UX Test (not tested real user yet)	[10]

Based on the literature analysis, none of the studies on the mobile flight and hotel booking application has applied two combined usability evaluation methods (UEMs) which are heuristic and UX test. The majority of the studies applied inquiry methods by distributed questionnaires and interviews [5], [6], [11], [7]–[9]. Administering questionnaire and interview are good to be conducted on mobile application as it is quick in getting the sample, but the questionnaire is often measured by user preference, not application usability [12]. In addition, for unstructured interview method, it will find difficulty in interpreting result [12]. Whereas this study will involve usability experts and real users. Holzinger conducted a study which showed that usability inspection methods need to be combined with user test method such as heuristic evaluation (HE) or cognitive walkthrough (CW) must be combined with direct user test [24]. The much different issues found by experts and actual users, the less reliable are the results.

III. METHODOLOGY

In general, the study of this research applies quantitative methodology during the data collection phase. As this study is an experimental research, the observations are conducted during UX test in order to collect data for usability metrics measurement of effectiveness and efficiency. The data is analysed to find the effectiveness and efficiency percentage of the application. In addition, the data collection for satisfaction metric is gathered by administering satisfaction questionnaire.

Firstly, this research is conducted by reviewing articles that are related to usability evaluation methods (UEMs). During literature review phase, many kinds of usability evaluation methods (UEMs) are in practice. This phase is aimed to explore all the methods available in the usability study. The second phase is followed by heuristic evaluation. This usability evaluation is conducted by usability experts to evaluate the mobile application by obeying heuristic evaluation principles of the touchscreen-based mobile device [20].

All the issues found by experts will be combined and analysed in order to produce heuristic evaluation report [25]. Furthermore, the third phase is followed by UX test that is often conducted in ordinary UX test by 5 to 10 users per test round [22]. Before conducting UX test, the test scenario is created based on the issues found in the heuristic evaluation report. In addition, users also will be asked to fill demographic questionnaire before the test started. During the test, the user will be asked to perform the test scenario given accordingly. In the other hand, evaluators will evaluate the user's performance by observing and analysing what user's doing. The fourth phase is followed by measuring usability metrics. There are three usability metrics [14] will be measured. The effectiveness and efficiency metrics are measured during users perform the test. Meanwhile,

satisfaction metric is measured by administering satisfaction questionnaire. The questions and answers are structured by using a 5-point Likert scale that ranges from 1 to 5.

All the questions are designed in order to get feedback from users on how they feel about Ticket.com application that has been chosen as mobile flight and hotel booking application. It is a critical application that can represent the study of this usability testing. The application users are more than 1 million. With the Likert scale, the users or participants indicate how they strongly agree or not at all in using this mobile application with “strongly disagree” equating to “1” and “strongly agree” equating to “5”. By having three usability metrics which are effectiveness, efficiency and satisfaction, the usability score of the application can be calculated in percentage.

The fifth phase of this study is to propose the recommendations and solutions based on the observation that has been conducted during UX test. In this phase, the issues found will be elaborated in details, and finally provide the recommendation to improve the application. Lastly, the sixth phase of this study will discuss and conclude all the findings found during this research. Research design diagram for this study is depicted in Figure 1 below:

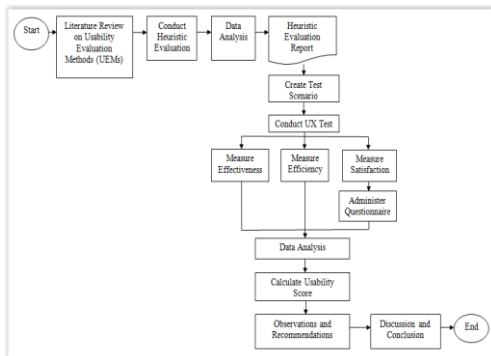


Figure 1: Research design diagram

The data collection methods for this study are conducted by observation and questionnaire. The data gathered during observation in UX test is analysed to measure usability metrics of effectiveness and efficiency. Both metrics are measured by the successful completion of criteria breakdowns from test scenarios [26] by analysing users' success rate. If the user successfully completes the task, it will be marked as “Yes”. This mark is given the full credit of 100%. For task criteria that are not matched will be marked as “No” and the credit is 0%. For this unsuccessful task criteria can be defined when users give up to complete the task or the users complete the task incorrectly. Also, there is 50% credit that is marked as “Partial” by evaluators in order to determine whether the mistake done by the user should be given partial credit or no mark at all [27]. Furthermore, the collection data of satisfaction metric is gathered by administering satisfaction questionnaire after the participants of test completed the test scenario given. The questions and answers are structured by using a 5-points Likert scale that ranges from 1 to 5 with “strongly disagree” equating to “1”

and “strongly agree” equating to “5”.

The participants of UX test will be selected from novice users with no previous experience on the application or have experience in other mobile flight booking application. The sample of users will be selected randomly with the minimum of five to ten users as increasing the number of users will not make any differences in the result [11], [23]. Meanwhile, for heuristic evaluation, the evaluators are chosen from people who have worked as usability engineer or has knowledge on this area. The minimum of evaluators is two in order to find the issues on the application.

The data collected to measure the percentage of effectiveness, efficiency and satisfaction are adopted to the equation that is proposed by [28]. Microsoft Excel will be used to analyse the data and present the data in the bar chart. Another software that will be used is WonderShare MirrorGo in order to the remote mobile phone to PC during UX test.

For this study on the mobile flight and hotel booking application (Ticket.com), the findings are analysed based on the observation when users are completing test scenario during UX test and questionnaire after completed the test. The analysis of the data gathered is calculated to get the effectiveness, efficiency and satisfaction in percentage. The average of these three scores is calculated to achieved usability score of the application in percentage with a number between 1 and 100. As a result, the recommendations and solutions are proposed based on the issues found during the observation.

IV. RESULT AND ANALYSIS

This section presents the results and findings of usability testing on mobile Ticket.com application. Heuristic evaluation was conducted by 3 usability experts to find issues on the application by following heuristic evaluation principles. The heuristic evaluation results were used to create the test scenario of UX test. Meanwhile, UX test was conducted by 10 users that were chosen randomly from both novice users with no previous experience and users with experience on mobile booking application. The UX test was conducted in a room or laboratory-like area that remote the mobile phone to PC, involved observation on the test participants while they performed several tasks given. The metrics measured on the application are effectiveness, efficiency and satisfaction. The test results about application revealed several strengths and weaknesses on the application as detailed in this section.

A. Heuristic Evaluation Result

Heuristic evaluation is evaluated individually by each usability experts in order to find the issues on the application. All the issues found by each usability experts will be combined to produce the heuristic evaluation report. The following table briefly reviews the most significant usability issues found by 3 usability experts and the recommendations for addressing them.

Table 2
Heuristic Evaluation Results

No	Issues	Heuristics	Severity	Recommendations
1	The form and checkbox of return flight will make users mistaken on the purchasing flight ticket.	TMD5: Error prevention	Major	<ul style="list-style-type: none"> • Provide checkbox of return separately with the return form. • Once user unchecks the checkbox of return, the return form should be blurred. • The application should keep informing users about all process of booking steps.
2	The users do not know in which step they are in the application.	TMD1: Visibility and system status	Major	<ul style="list-style-type: none"> • The user should be able to know in which step they are now, how many more steps to complete the booking process.
3	The main search function in hotel displays the different result with search function provided in the list of hotel name. Sometimes cannot find any lists of the hotel.	TMD8: Efficiency of use and performance	Major	All the hotel search function should be able to load and display the same result.
4	- Cannot exit from the application. - The user needs to press "home button" on the device to exit the application.	TMD3: User control and freedom	Major	When the user presses the back button on device more than twice on the home screen, the users should be able to exit from application.
5	The language of the error message is not similar to the user preference language chosen.	TMD4: Consistency and standards	Major	The application should provide the similar language of error message with the language preference to prevent user misunderstanding.
6	Some of the payment methods do not function.	TMD8: Efficiency of use and performance	Major	<ul style="list-style-type: none"> • All the payment method listed should function well. • Recommended providing only payment method available in the country of user location.
7	There is no main menu on every screen in booking steps in order for user to change language preference and currency	TMD8: Efficiency of use and performance	Major	The application should provide the main menu on every screen to make easier for the user to change language preference and currency.
8	The label name of the checkbox is different with the error message displayed. In the label, it uses "infant" but the error message use "baby"	TMD4: Consistency and standards	Minor	Use the same term to provide language consistency
9	There is no back button when the user received "no available room in the hotel".	TMD3: User control and freedom	Minor	The application should provide the back button to allow users choose another available hotel.
10	No error message or alert is given to the user when selecting the wrong number of the infant.	TMD10: Help users recognise, diagnose, and recover from errors	Minor	Provide the appropriate message to inform users when the wrong input.
11	There is no error message when users would like to save the profile. The user did not recognise the errors.	TMD10 : Help users recognise, diagnose, and recover from errors	Minor	When save button in add profile does not work, the error message should appear to inform users.
12	- Users are not informed which data field is required to fill in passenger information. - No highlight on the mandatory field when users do not fill the information.	TMD1: Visibility of system status	Minor	<ul style="list-style-type: none"> • Recommended putting the mandatory field indicator below the text box. • Highlight the mandatory field to inform users to fill in.
13	No error message on password submission when users input the wrong password in the login page.	TMD10 : Help users recognise, diagnose, and recover from errors	Minor	Provide error message to inform users.
14	Availability of the language is limited.	TMD2: Match between system and the real world	Minor	Provide language based on the available currency or user location.
15	Catagories filter on hotel search function. Example: when the user searches a country name such as Malaysia, provide the region of the country result to be selected in order to lists all the hotel name.	TMD8: Efficiency of use and performance	Minor	Should filter the search result accordingly.
16	No delete/remove button in some of expired booking transaction history.	TMD12: Physical interaction and ergonomics	Minor	The application should provide the delete/remove button for all expired booking history.

Based on the table above, the experts have found 7 major issues and 9 minor issues. The evaluation process was conducted by following the 12 heuristics principles for the touchscreen-based mobile device [20]. The severity of the issues is decided by following the Nielsen's severity ranking scale (SRS). There are 4 severity scales proposed by Jakob Nielsen such as *usability catastrophe* where the issue is imperative to fix before the product can be released. The second scale is *major* usability problem is described as the high priority to be fixed. Meanwhile, the *minor* usability problems, fixing the issues should be given low priority. Lastly, *the cosmetic problem* needs to be fixed if there are extra time available on the project. Thus, the

recommendations are proposed to help the developers to fix the issues on the application.

Furthermore, the issues found on the heuristic evaluation will be used to designed test scenario on the UX test. The set of the test scenario is created as much as possible to include all the main function in flight and hotel booking process.

B. UX Test Scenario Design

Prior to test execution, test scenario is designed in order to test the application. Task #1 is designed to allow user login to the application in order to process flight and hotel booking. Task #2 is designed to test the flight booking process in the application. But, in booking date for task #2, #3 and #5 is

filled differently as all participants log in to Tiket.com use the same email provided in tasks description. Task #3 is designed to test the hotel booking process and task #4 is designed to test the transaction history of the application. Task #5 is designed to see how user learned from their mistakes and errors in task #2 where the users are asked to book the hotel with different currency. This task will show how the application easy to learn and recover from error. Task #6 is designed to test the update profile menu and task #7 is to log out of the application. Task #8 is to ask user exit on the application. It shows that there is no exit function on the application as the error has been found during heuristic evaluation.

Table 3
UX Test Scenario Design

Scenario	Tasks Description
Task #1	Log in Tiket.com as credential below: E-mail: ma.suay@gmail.com Password: ma1212m
Task #2	Book one-way flight From: Kuala Lumpur (KUL) To: Banda Aceh (BTJ) Date: __ April 2017 Currency: Dollar Singapore Payment Method: ATM Transfer
Task #3	Book a room hotel Location: Banda Aceh Check-in: __ April 2017 and Check out: __ April 2017 Guest: 2 Adult Payment Method: ATM Transfer
Task #4	View transaction history on application
Task #5	Book a room hotel Location: Kuala Lumpur Check-in: __ April 2017 and Check out: __ April 2017 Guest: 2 Adult Currency: MYR Ringgit Malaysia Payment Method: Cimb clicks
Task #6	Update new profile with your credential
Task #7	Log out from Tiket.com
Task #8	Exit from application

C. Effectiveness Measurement

The metric of effectiveness is measured by doing the observation on users during test execution. The task criteria that were being observed in every task in test scenario such as "user understand the flow how to do the task", "user does not have problem to take the right steps to do the task", "user do not need assistance to complete the task", and "user succeed to complete the task". Every successful task criteria will be marked "yes" for 100% credit, "partial" is 50% for partially successful and "no" is 0% for the unsuccessful task.

Based on the observation and findings, we can conclude that the flow in flight booking is not ease-to-use as the result shows that no users successfully complete the task #2 (Book one-way flight). Some users are able to complete the payment process, but the currency used is not changed as the task requested. For this case, the users will be assigned "partial" rather than unsuccessfully complete the task. The partial mark is 50% credit. For other users that encountered internal server errors and problem in passenger information page will be assigned 0% credit or unsuccessfully complete.

As regard task 1, there is only user 7 that unable to complete the task. The problem is caused by the login button that did not work as expected even the user already input correctly. In the observation of task #3, there are 4 users unable to complete the task which are user 1, user 2, user 3 and user 8. User 1 encountered network application issue,

user 2 encounters internal server error, user 3 encountered unavailable atm transfer payment method and user 8 decided to give up on the task. In addition, 90% of task #4 were successfully completed. There is only one user is unsuccessfully complete the task and take so much time.

The task #5 is intentionally created to see how users learn from their mistake on task #2. Based on the observation, only 4 users are successfully complete the task #5. 2 users are partially complete and 4 users are unsuccessfully complete the task. This results can be concluded that the application is not easy to learn. In addition, there is one user who is partially complete task #6 (update the new profile with your credential). The partial credit is given as the users understand how to do the task, but have internal server error that makes them cannot successfully complete the task (user 1). Another 2 users are unsuccessfully complete the task. User 2 and user 8 encountered the same problem in saving profile where there was no error message and users cannot recognise the error. In regard task #7, there is no issue found where the task 100% successfully completed.

In the results of observation task #8, Most of the users (90%) press home button on the device to exit from the application where the back button does not function as the normally way to exit. Some of the users looked for physical exit button on the application, but then finally press the home button to exit. There is only one user that looked for the physical button to exit and did not press the home button. In another word, we could say there are no users able to exit from the application as user expected.

The effectiveness results of users success rate evaluation can be analysed to a simpler effectiveness analysis that is described as in the table below:

Table 4
Effectiveness Analysis

	User 1	User 2	User 3	User 4	User 5	User 6
Yes	12	23	25	19	17	25
Partial	6	3	1	1	3	4
No	14	6	6	12	12	3
	User 7	User 8	User 9	User 10	Subtotal	
Yes	19	13	23	25	201	
Partial	6	7	5	3	39	
No	7	12	4	4	80	
TOTAL						320

According to the table above, it shows that there are 32 tasks criteria with 10 attempts per task and 320 of the total attempts. 201 attempts of task criteria were successfully completed and 39 task criteria were partially successful. However, there are 80 of task criteria that were unsuccessful completed which will be calculated as $80 \times 0\% = 0$.

D. Efficiency Measurement

The metrics of efficiency is also measured by doing the observation on users during test execution. The task criteria that were being observed in every task in test scenario such as "user select the right steps at the first try", "user easily recover from errors", "error and mistake did by the user are minimal", and "user does not take much time to complete the task". Every successful task criteria will be marked "yes" for 100% credit, "partial" is 50% for partially successful and "no" is 0% for the unsuccessful task.

Based on the observation, there are only 2 participants who select the right step at the first try in flight booking. In

addition, 8 out of 10 users take much time to complete task #2. In another word, we could say the flight booking process is not efficient. Another issue is on task #5, there are only 3 participants who are able to select the menu at the first try, but only one participant who do not take much time to complete the task. As the result indicates that the application is not easy to learn as the purpose of this task was designed to see how participants learn from their mistakes in task #2.

The efficiency results of user success rate evaluation can be analysed to a simpler efficiency analysis that is described as in Table 5 below:

Table 5
Efficiency Analysis

	User 1	User 2	User 3	User 4	User 5	User 6
Yes	14	21	25	19	16	26
Partial	1	2	0	0	1	2
No	17	9	7	13	15	4
	User 7	User 8	User 9	User 10	Subtotal	
Yes	18	15	23	27	204	
Partial	4	2	4	2	18	
No	10	15	5	3	98	
TOTAL					320	

According to the table above, it shows that there are 32 task criteria with 10 attempts per task and 320 of the total attempts. 204 attempts of task criteria were successfully completed and 18 task criteria were partially successful. However, there are 98 of task criteria that were unsuccessful completed which will be calculated as $98 \times 0\% = 0$.

E. Satisfaction Measurement

The metric In order to measure the satisfaction of the application, the post-questionnaire was administered to participants. The participants need to answer all questions after performing all test scenario that is designed by using a 5 point Likert scale.

All questions on the satisfaction questionnaire were designed to measure the satisfaction of users in using the Tiket.com application. The design of satisfaction questionnaire is adapted from system usability scale (SUS) questionnaire that was invented by John Brooke in 1986 [30]. It is also administered to know how participants feel about the application, do they like to use it and does the application is easy to use or not.

Using the numerical value of 5 points Likert scale with “strongly disagree” equaling to “1” and “strongly agree” equaling to “5”, each question answered by 10 participants offers a possible positive response factor of 60 points. Thus, there is 600 points or 100% satisfaction for 12 questions. In order to get the satisfaction rating for the Tiket.com, the following equation is used:

$$Satisfaction(\%) = \frac{Answer}{Total} \times 100 \tag{1}$$

Based on the Equation (1), UX testing with 10 participants has shown the satisfaction rating for Tiket.com is approximately 61.67%. All the users that have high satisfaction score are the users that have never bought the ticket on their mobile. In another statement, we could say, they have never experienced how a good usability of others mobile booking application. In contrast, the users that have previous experience purchasing flight tickets or hotel booking on their mobile (user 6 & 7), the result of satisfaction is very low. In conclusion, the users feel the application is not easy to use compared to their favourite mobile booking application.

F. Usability Score

As the three metrics (effectiveness, efficiency and satisfaction) have been calculated in percentage, the usability score can be calculated by averaging these three scores to define the usability of Tiket.com with a number between 1 and 100. Usability Score for Tiket.com from all users can be derived from the following equation:

$$Usability(\%) = \frac{Effectiveness + Efficiency + Satisfaction}{3} \tag{2}$$

From the Equation (2), the UX testing with 10 users has shown the usability level for Tiket.com is approximately 65.71%. According to Sauro (2011) the average usability score is 68%. Thus, the usability of Tiket.com is below average or C- in the letter grade [29]. In another more specific statement, we could say there are probably serious problems on the application usability that need to address.

As a matter of fact, the issues found on UX testing can be fixed in order to have a user-friendly mobile booking application. Furthermore, the analysis of these usability metrics which are effectiveness, efficiency and satisfaction [14] defined how effective, efficient and satisfy the Tiket.com is. Furthermore, in order to attract more users to use the application, the major issues found should be considered as the high priority to be fixed as the satisfaction result of the application is low (61.67%). Therefore, this usability result can be used by Tiket.com usability engineer in order to improve the usability of the mobile application.

G. Observation and Recommendation

The observation was conducted while the user was completing the task scenarios on UX test. All the errors made by users and application errors were noted. The details explanation of errors or issues found during observation is described as below:

Table 6
Observations and Recommendations

Issues	Observations	Recommendations
No error message in login page.	When user input wrong password in the login page, there is no error message to tell users. The users keep pressing the login button but it didn't work. Some of the users asked help from evaluator what to do, some of them try to discover the error by retyping the e-mail and password on the textbox.	Provide the appropriate message to inform users when wrongly input.
Login button in login page is sometimes not working.	The login button is not working when user 7 try to log in the application. The user has entered the correct e-mail and password several times, but the button still not working. Then as user continued to complete task #2 where it needs to log in the application, then the user was able to login to the application successfully.	The main function such as login should work as user expected.

Issues	Observations	Recommendations
Checkbox of return flight.	The checkbox of return flight makes user mistaken in booking one-way flight as the default of checkbox is checked. Based on the observation, almost all users facing this problem. They are all get confused on checkbox provided in return flight form. Some of them do not notice the checkbox of return. As the results, in task #2 users were asked to book the one-way flight, but some users booked the two-way flight by choosing the same date of the return flight. There are only two users are partially success in completed the task #2 that will be assigned 50% credit.	<ul style="list-style-type: none"> • Provide checkbox of return separately with the return form • Once user unchecks of the checkbox, the return form is recommended to be blurred.
Most of the users have problem to look for currency menu as it is a part of test scenario #2 and #5.	Most of the users take very long time in booking flight (test scenario #2). They have the problem to find the currency menu. Some users, use back button to go to home page and find the currency menu after search result came out. Seems users expect the destination and date of flight chosen are not deleted yet. In fact, after choosing the currency menu, the user has to start flight searching again. Some users expected at the end of ordering process they will find the currency menu and try clicking on the amount as they expected the currency menu would pop up. Some of the users gave up and continue to next test scenario.	The application should provide the currency menu in every the page of the booking process.
Users do not know what step they are in the booking process.	Users want to know how many steps more to complete the booking process. Some of the users got confused if they are already completed the task or not.	<ul style="list-style-type: none"> • The application should keep informing users about all process of booking steps. • The user should be able to know in which step they are now, how many more steps to complete the booking process.
Errors in passenger information in flight and hotel booking	Based on the observation, there are 4 errors found on passenger information page. The errors are explained in the list below: <ul style="list-style-type: none"> • There is no error message on passenger information page when users incomplete fills in data. Users hard to recognise errors made. • Users are not informed which data field is required to fill in passenger information. No highlight on the mandatory field when users do not fill the information. • The text box of the name in passenger information does not allow to have more than 20 characters as their full name as in passport is more than 20 characters. • Error in baggage check-in (user 6, user 7 and user 9). Based on the observation, the error happened when users choose AirAsia flight without baggage check-in. Users already fill in all information, but the application keeps inform users to choose check-in baggage. So most of the users cannot successfully complete the task #2. There is only one user (User #6) that recognise this error and choose another airline. But this user is also not able to complete the task because during booking process there was the server error. 	<ul style="list-style-type: none"> • Provide the appropriate error message to inform users to fill in the incomplete data. • Recommended putting the mandatory field indicator below the label. • Highlight the mandatory field to inform users the to fill in mandatory data. • Allow more than 20 characters in the textbox of the name in passenger information. • The check-in baggage error should be fixed in order to proceed flight booking.
Error in search result of hotel.	Sometimes there is no search result found in the list. The users need to search several times. There are many kinds of error message pop up such as "internet connection problem" where actually the internet connection is stable. It seems the errors caused by the network application. Another error message is "Ups, something wrong happen". One of the users use the second search in hotel booking, but there is no result found.	<ul style="list-style-type: none"> • The search function of the hotel should work properly as user expected. • All the hotel search function should be able to load and display the same results.
Error in payment method	Some of the available payment methods are not working. As the test scenario #2 and #3 to complete payment by "ATM transfer", some users found there is no that kind of payment method available. Some user found internal server error message in the payment process.	<ul style="list-style-type: none"> • All the payment method listed should function well. • Recommended to provide only payment method available in the country of user location.
Error in updating profile.	When users save the update profile, sometimes it takes too long and comes out with an error message "internal server error".	The save button in update profile should function well.
Back button on home page does not work as an exit in application.	Some users press the home button to exit from the application. Some users after press back button several times, then try to find the physical exit button. Some users just give up after searching the exit button. As the results, there is no one able to complete task #8.	<ul style="list-style-type: none"> • Provide function of double click on the back button on the device should exit from the application and also sending the message like pressing one more time will exit from the application. • It is recommended to have physical exit button.

V. DISCUSSION

The main findings of this study revealed that the users' expectancy is the complex matters that may happen in any kind of software testing. Most of the issues found in heuristic are also found in UX test. In another word, we could say these two usability evaluation methods (UEMs) is a good combination to evaluate mobile application. This result substantiates to Holzinger's study on combining any usability inspection methods with direct user test [24].

The three metrics such as effectiveness, efficiency and

satisfaction are strongly matched in this usability measurement. It means the observation and questionnaire results are strongly interrelated. However, the study indicates the user's perception toward application may show differently with satisfaction results, most of the participants answered neutral or agree to the first post-test questionnaire of "I would use this application again". In contrast, the satisfaction results have shown the score is below average (62%). These results indicate the users are not satisfied with the application.

Another issue is regarding users' understanding about test scenario. During observation, we found that some users are

very confident in completing the tasks in UX test. The users feel they completed the task correctly, but actually they failed in completing it. In addition, we have observed in flight booking process, users tick check, then uncheck in the return of flight several times. The task is very clear that asked users to book a one-way flight, but users get confused in the return checkbox. These results indicate the users do not really understand the task given. It needs more briefing prior to the test execution to make sure users understand what they are going to do. Furthermore, the test scenario design should be tested many times before the test execution.

Overall, the result of usability testing of the application has shown the weakness of the application which is below average (66%). It indicates the application probably has serious problems that need to be fixed. More test with different operating system need it to validate and generalise these results. Thus, the evaluation of different operating system may present the comparative outcome.

In conducting this usability testing, it has limitation such as remote software (MirrorGo) is not able to record screen activities after one-minute recording. It also found other remote software is not compatible with android version 4.2.2 because we tested the application with OPPO N1 on this version. Furthermore, the test is not conducted in formal usability laboratory with limited usability tools. However, the result of this study can be adopted by the developers and all usability practitioners to deliver a user-friendly mobile application that leads to high customer satisfaction and also revenue improvement.

VI. CONCLUSION AND FUTURE WORK

This paper has presented the main challenging issues that are faced by mobile booking application after conducting heuristic and UX test. As regard to the objectives and results of this study, it can be concluded that the usability score of Tiket.com is below average (66%) where there are also a lot of issues found during the heuristic evaluation and UX test. Most of the issues found by expert are also found during UX test. The result reveals that the application is not usable that need to be improved in order to have a good mobile booking application.

Overall, the aims of this study are to evaluate the usability of mobile Tiket.com application in order to measure the effectiveness, efficiency and user satisfaction on the application. The current study on Tiket.com application illustrated that the mobile commerce application in the application market is not all easy to use and user-friendly. It proves that the application needs to be updated several times in order to fix the issues on the application. Thus, the current study on this application can be used to prevent errors and mistakes that will possibly encounter by users.

Practically, this study contributes by proposing the solutions and recommendation for Tiket.com in order to improve the usability application. The recommendations are presented accordingly to the issues found. Overall, this study is significant to all mobile booking or m-commerce application developers to develop a good mobile application in order to achieve competitive advantages. It is also important to have the mobile application that is user-friendly and easy-to-use.

However, this issue required further investigation by conducting think aloud method in formal usability laboratory which has the one-way mirror to an observation room. It is

also highly recommended to have usability tools such as eye tracker glasses and camera recorder that records user's behaviours and face expressions.

REFERENCES

- [1] Nielson, "What's in-store for online grocery shopping," 2017. [Online]. Available: <http://www.nielson.com/us/en/insights/reports/2017/whats-in-store-for-online-grocery-shopping.html>. [Accessed: 15-Feb-2017].
- [2] A. Hussain, N. L. Hashim, N. Nordin, and H. Mohamad Tahir, "A Metric-Based Evaluation Model For Applications On Mobile Phones," *J. ICT*, pp. 55–71, 2013.
- [3] J. Nielsen, "MOBILE ACTS AS A NERVE CENTER OF THE CANADIAN SHOPPING EXPERIENCE," 2016. [Online]. Available: <http://www.nielson.com/us/en/insights/news/2017/mobile-acts-as-a-nerve-center-of-the-canadian-shopping-experience.html>. [Accessed: 15-Mar-2017].
- [4] Bronto Software Inc., "Younger consumers turn increasingly to mobile devices for buying as well as browsing," 2016. [Online]. Available: <https://www.digitalcommerce360.com/2016/04/25/younger-consumers-turn-increasingly-mobile-devices-buying/>. [Accessed: 14-Mar-2017].
- [5] R. Molich and C. Wilson, "Tips and Tricks for Avoiding Common Problems in Usability Test Facilitation," *Chi 2008*, pp. 2379–2382, 2008.
- [6] X. Sun and A. May, "The role of spatial contextual factors in mobile personalization at large sports events," *Pers. Ubiquitous Comput.*, vol. 13, no. 4, pp. 293–302, 2009.
- [7] A. Moller, L. Roalter, S. Diewald, J. Scherr, M. Kranz, N. Hammerla, P. Olivier, and T. Plotz, "GymSkill: A personal trainer for physical exercises," *2012 IEEE Int. Conf. Pervasive Comput. Commun. PerCom 2012*, no. March, pp. 213–220, 2012.
- [8] U. Knebel, J. M. Leimeister, and H. Krcmar, "Personal mobile sports companion: Design and evaluation of it-supported product-servicebundles in the sports industry," *Proc. 15th Eur. Conf. Inf. Syst. ECIS 2007*, pp. 81–92, 2007.
- [9] F. Buttussi, L. Chittaro, and D. Nadalutti, "Bringing mobile guides and fitness activities together: a solution based on an embodied virtual trainer," *Work*, vol. Espoo Finl, pp. 29–36, 2006.
- [10] I. Burmistrov, "Mobile air ticket booking," *VTT Symp. (Valtion Tek. Tutkimuskeskus)*, no. 258, pp. 115–118, 2009.
- [11] F. Gunduz and A. S. K. Pathan, "Usability improvements for touchscreen mobile flight booking application: A case study," *Proc. - 2012 Int. Conf. Adv. Comput. Sci. Appl. Technol. ACSAT 2012*, pp. 49–54, 2012.
- [12] P. Tsai, "a Survey of Empirical Usability Evaluation Methods," 2006.
- [13] M. Kutar and A. Hussain, "Usability Metric Framework for Mobile Phone Application Usability Metric Framework for Mobile Phone Application," *ResearchGate*, no. June, 2009.
- [14] ISO, "Ergonomic Requirements for Office Work with Visual Display Terminals (VDTs)—Part II Guidance on Usability," 9241–11, 1998.
- [15] M. C. Trivedi and M. A. Khanum, "Role of context in usability evaluations: A review," *Adv. Comput. An Int. J.*, vol. 3, no. 2, pp. 69–78, 2012.
- [16] B. Battleson, A. Booth, and J. Weintrop, "Usability testing of an academic library Web site: a case study," *J. Acad. Librariansh.*, vol. 27, no. 3, pp. 188–198, 2001.
- [17] D. Pinelle, N. Wong, and T. Stach, "Heuristic evaluation for games: usability principles for video game design," *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, pp. 1453–1462, 2008.
- [18] E. Bertini, S. Gabrielli, and S. Kimani, "Appropriating and assessing heuristics for mobile computing," *AVI '06 Proc. Work. Conf. Adv. Vis. interfaces*, pp. 119–126, 2006.
- [19] A. S. Hashim, W. A. Wan Ahmad, and R. Ahmad, "Usability and Effectiveness of Mobile Learning Course Content Application as a Revision Tool," *Comput. Technol. Appl.*, vol. 2, no. June, pp. 148–157, 2011.
- [20] R. Inostroza, C. Rusu, S. Roncagliolo, C. Jiménez, and V. Rusu, "Usability heuristics for touchscreen-based mobile devices," *Proc. 9th Int. Conf. Inf. Technol. ITNG 2013*, no. 2241, pp. 662–667, 2013.
- [21] ISO, "Ergonomics of human-system interaction," 9241–210, 2008.
- [22] A. Kaikkonen, A. Kekalainen, M. Cankar, T. Kallio, and A. Kankainen, "Usability testing of mobile applications: A comparison between laboratory and field testing," *J. Usability Stud.*, vol. 1, no. 1, pp. 4–16, 2005.
- [23] F. Gündüz and A. S. K. Pathan, "On the key factors of usability in small-sized mobile touch-screen application," *Int. J. Multimed. Ubiquitous Eng.*, vol. 8, no. 3, pp. 115–138, 2013.

- [24] A. Holzinger, "Usability engineering methods for software developers," *ACM*, vol. 48, no. 1, pp. 71–74, 2005.
- [25] J. Nielsen, "Enhancing the explanatory power of usability heuristics," *CHI '94 Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, pp. 152–158, 1994.
- [26] J. Nielsen, "First Rule of Usability? Don't Listen to Users," 2001. [Online]. Available: <https://www.nngroup.com/articles/first-rule-of-usability-dont-listen-to-users/>. [Accessed: 20-Mar-2017].
- [27] M. Ismail, N. M. Diah, S. Ahmad, N. A. M. Kamal, and M. K. M. Dahari, "Measuring usability of educational computer games based on the user success rate," *SHUSER 2011 - 2011 Int. Symp. Humanit. Sci. Eng. Res.*, pp. 56–60, 2011.
- [28] ISO/IEC, "Software Engineering-Product Quality," 9126–4, 2004.
- [29] A. Bangor, P. Kortum, and J. Miller, "Determining What Individual SUS Scores Mean : Adding an Adjective Rating Scale," vol. 4, no. 3, pp. 114–123, 2009.
- [30] J. Sauro, "Measuring Usability with The System Usability Scale (SUS)," 2011. [Online]. Available: <https://measuringu.com/sus/>. [Accessed: 01-May-2017].