International Journal of Advanced Science and Technology
Vol. 29, No. 8, (2020), pp. 180-187

The Effectiveness, Efficiency and Reliability-in-Use of Daylio Mobile App

Azham Hussain[1], Emmanuel O.C. Mkpojiogu[1,2], Idyawati Hussein[1*], Omer Mahdi Muhi[1], Muhammad Hibatul Hakimi Yosri[1]

[1]*School of Computing, Universiti Utara Malaysia, 06010 UUM, Sintok, Malaysia*
[2]*Department of Computer and Information Technology, Veritas University, Abuja, Nigeria*
[1*]*idyawati@uum.edu.my*

Abstract

This study conscripted 14 participants to examine the effectiveness, efficiency and reliability-in-use of a mood tracker app. A usability laboratory protocol was utilized to elicit task completion (success) rate, task completion time, and task error metrics with which to measure effectiveness, efficiency and reliability-in-use of the Daylio app respectively. The researchers conducted the evaluation in a lab setting. Overall, the outcomes reveal that the Daylio mobile app is effective, efficient and reliable in use. However, a close scrutiny indicated that there are pressing issues that need to be fixed to further enhance the quality of the app. Fixing these issues will certainly improve the success and performance of the app, its market penetration, as well as prompting good users/customers loyalty and retention.

Keywords: *App, effectiveness, efficiency, mood tracker, reliability in use*

1. Introduction

This study focuses on the evaluation of the effectiveness, efficiency, and reliability-in-use of Daylio: a mood tracking app that handles the monitoring stressful conditions of everyday life. Daylio was created by Relaxio S.R.O. and is both on Android and iOS platforms. The app is a micro-diary based app that enables the recording of users' current mood without the need of them typing any words. The app's design is simple. It also displays the mood history in a graph format to assist users keep track of their mood. Effectiveness, efficiency and reliability in use are the three main thrust in usability testing. Effectiveness is measured using task completion. Efficiency is measured using time on task while reliability in use is captured using error rates. There are various guidelines for testing the usability of desktop apps. Nonetheless, those established concepts, methodologies, and approaches commonly used in traditional human–computer interaction research are not always applicable to mobile applications due to the mobility and other distinct features of mobile devices and wireless networks. Ideally, usability testing of mobile applications should be carefully designed to cover all or most possible situations of a mobile environment, in reality; however, this poses many challenges. For example, it is difficult to foresee the exact situations of the application use—users may be standing, walking, or sitting in a dark or bright environment while using an application. As a result, a usability test may have to concentrate only on certain aspects of a mobile application and sacrifice others. Furthermore, traditional research methodologies used in usability testing, including controlled laboratory experiments and field studies, have various limitations in a mobile environment, such as ignoring the mobile context or lack of sufficient procedural control [1-3][16-20][27].

In remote usability evaluation, evaluators and users are separated in space and possibly time during the evaluation. This type of evaluation is becoming increasingly important for a number of advantages it offers. Indeed, it allows the collection of detailed information on actual user behavior in real contexts of use, which is especially useful in contexts in which it is not possible (or convenient) having an evaluator directly observing or recording the session. In addition, the fact that the users carry out the evaluation in their familiar environments contributes to gain more 'natural' users' behavior. In order to have a complete

ISSN: 2005-4238 IJAST

picture of what users did during the session and derive consequent conclusions about the usability of the application, it is crucial for the evaluators to reconstruct not only the interactions that users carried out during the session, but also the contextual conditions that might have affected the user interaction itself. Indeed, if such conditions are not completely known, the evaluators might draw incorrect conclusions about the usability of the considered application. This problem becomes even more difficult to address when dealing with mobile applications. Indeed, while for desktop applications lack of co-presence between users and evaluators can be compensated to some extent by equipping the test environment with devices such as web cams. Mobile applications require different solutions that are able to flexibly support evaluation in different contexts without being too obtrusive on the user side. When dealing with remote applications for mobile devices, there are some additional problems that make it more difficult to gather data to remotely evaluate such applications. For instance, we have to take into account the more limited capability of mobile devices, which imposes constraints on the kinds of techniques to be used for tracking user interactions. In addition, there is the further problem of detecting the environmental conditions in which the session takes place [5][7-9][11-15][19][21-25].

There are many methods for usability evaluation. These methods can be categorized into four main areas, namely, heuristic evaluations, cognitive walkthroughs, usability testing, and comparison against guidelines. To decide which evaluation method to use depends on the strengths and weaknesses of the method, as well as its applicability with regards to the researcher's objectives. There are increasing concerns and critiques of the validity and effectiveness of those methods from human factors and human-computer interaction aspects [4]. Usability tests, being part of the usability engineering process, are performance measurement to determine whether usability goals are achieved. It is usually conducted in laboratories with test participant(s) performing a pre-defined set of tasks while data on performance measures are documented. Usability tests are conducted to collect quantitative data on usability problems (difficulties that users come across while using the application or the device), performance of devices (in terms of accuracy, time to carry out functions, etc.) and mental/physical demands of using device, etc. Many types of usability problems will arise from the tests are conducted and can be generalized according to their importance. Typical categorization include: i) critical problems, then ii) serious problems to the lesser iii) cosmetic problems. Statistical analysis can then be done for comparison between tests. With such measures, the design of the evaluated device or system and its interface can be improved based on the result of evaluation [6].

The usability of mobile gadgets and their applications varies from other PC systems, on the grounds that their qualities are different. The software needs of handhelds, for example, PDAs and mobile phones, influence the development procedure of mobile applications, as these are inserted in the phones during production or installed by clients from several mobile software dissemination platforms, for example, Apple's App Store and Google's Android Market. Users will in general pick mobile applications that are easy to learn, allow them to perform specific task in lesser time, and looks more user-friendly since they are less computer-oriented. Previously, the usability of software systems was tested subjectively, and the procedure was not very much characterized. Researchers would choose the parts of usability to test and measure what they thought are vital. At the same time, usability estimation and analysis methods and methodologies were being created. Lab experiments, field studies, and hands-on estimation are some of techniques frequently used by researchers [10][21][26-29]. Usability is a multi-construct and multidimensional quality. Effectiveness, efficiency and reliability-in-use are part of the dimensions that measure usability. In this study, these qualities will be measured and evaluated.

Effectiveness is the degree at which a user is successful in carrying out his/her goals (tasks) while interacting with an application in a given context of use. Efficiency on the other hand is the degree at which a user accomplishes his/her goals during an interaction with a product with minimal resources in a specified context of use. Lastly, reliability in use is the extent to which a user accomplishes his/her goals while interacting with a product with minimal errors in a given context of use [26-30].

2. Methods

In this study, 14 participants comprising of 11 males and 3 females from the State of Kedah in Malaysia were conscripted into the study. The 14 participants included in the study include 7 postgraduate students and 7 civil servants. Of the 14 participants, 12 are novice users while 2 are moderate users in terms of experience. The following are the protocol employed for the test evaluation:

The Lab Setup: Lab method was employed for the usability testing of Daylio. Two labs in Universiti Utara Malaysia, UUM, were used: the Smart Room in the Library and the Inasis Tradewinds. The participants were given the freedom to choose the test location of their choice. There were two moderators involved in the evaluation, one in each of the two lab location. The participants were allowed to carry out the test at their convenience. Four tasks were given to participants to carry out.

Materials Used: Participants were given the permission to either use their personal phones or the moderator's phone. The moderator's phone is equipped with all applications to be used for the testing. Participants using their personal phones were instructed to install both the Daylio application and Mobizen screen recorder app on their phones before the commencement of the test. Each participant's test session were individually recorded for accuracy task time and error recording.

Description of Tasks: All participants were to use touch screens to carry out the test's 4 tasks. The facilitating moderator recorded the entire test sessions for all participants using screen recorder software. Each participant carried out his/her test individually, one participant at a time and also recorded individually (and in each case, the moderator observes). While the test lasts, the moderator collects data on task completion, time on tasks (task completion time) and error rates. The 4 tasks used in this test sessions are: i) Task 1: Check Status; ii) Task 2: Add New Mood Entry; iii) Task 3: Check Weekly Reports; iv) Task 4: Edit Activities. All participants were encouraged to think aloud while carrying out the respective tasks to conveniently elicit participants' reaction toward the app.

On the whole, the steps followed in the data collection and analysis process are as follows: Step 1: Selection of 14 participants and their briefing on test and task completion process. Step 2: Recording of each of the four tasks using a screen recording app. This is to enable the capturing of the task completion time and error rates. The time allocated to each participant to carry out all 4 tasks is 16 minutes, with 3 minutes given for each tasks. Step 3: On completion of the test sessions, the data collected will be analyzed using SPSS software.

3. Results

Task Completion time: Task completion time is a measure for the efficiency of the app. Table 1 displays the time taken by each user per task in comparison to the allowed 3 minutes for the task completion. In task 1, six users representing 42.86% of users completed their task in less than 1 minute. Seven users (50%) spend more than 1 minute but less than 2 minutes to complete the task. One participant however spent more than 2 minutes for this task. In task 2, ten participants (71.43%) spent less than one minute to accomplish task while the remaining 4 (28.57%) finished their tasks within 2 minutes. In task 3, eight participants (57.14%) completed the task within 1 minute while 4 (28.57%) completed theirs in less than 2 minutes but more than 1 minute. However, 2 participants (14.29%) spent more than 2 minutes in the task. In task 4, five participants (35.71%) completed their tasks within 1 minute while 8 (57.14%) completed theirs in less than 2 minutes but in more than 1 minute. However, 1 (7.14%) participant completed task in a time slightly above 2 minutes. On the whole, all participants completed all the tasks within the criterion time of 3 minutes. This implies that the application is fast and efficient for use.

ISSN: 2005-4238 IJAST

Users	Task1 (3 Minutes)	Task2 (3 Minutes)	Task3 (3 Minutes)	Task4 (3 Minutes)
1	1.49	1.45	1.26	0.44
2	0.26	1.15	0.22	0.52
3	1.17	0.40	0.28	1.00
4	0.18	1.20	0.45	0.40
5	1.20	1.25	0.20	2.03
6	0.30	0.21	0.23	0.41
7	1.10	0.41	2.02	0.35
8	0.55	0.37	1.55	1.43
9	0.35	0.42	0.57	1.23
10	1.36	0.56	2.07	1.32
11	1.55	0.54	0.54	1.03
12	0.36	0.39	1.03	1.23
13	2.02	0.39	1.03	1.23
14	1.20	0.50	0.37	1.03

TABLE 1. Participants' task completion time

Task Errors: Task error rate is a measure of the reliability in use. Errors are any actions that lead to a different destination other than the designated goal destination (that is, any wrong action taken by the user in the course of carrying out his/her tasks). In this study, most participants completed their respective tasks without errors. Per task, the maximum error made was 2 with the exception of one participant that had 5 errors in task 5. Generally speaking, the app is reliable in use. Table 2 provides details on the error data.

Users	Error Task 1	Error Task 2	Error Task 3	Error Task 4
1	1	0	0	1
2	0	1	0	0
3	1	0	0	0
4	0	1	2	0
5	2	2	0	1
6	0	0	0	0
7	0	0	5	0
8	0	0	1	0
9	0	0	1	0
10	1	1	3	2
11	1	1	0	0
12	0	0	0	0
13	2	0	0	0
14	0	1	0	0

TABLE 2. Participants' task errors

Task Completion (Success) Rate: Task completion or success rate is a measure of the effectiveness of an application in use. From Table 3, it can be seen that in all tasks, there was a 100% success rate. This means that the application is easy to understand, learn, operate and use and thus very effective as users were all able to accomplish their goals while using it.

ISSN: 2005-4238 IJAST

Participants	Task 1 (Check Status)	Task 2 (Add new entries)	Task 3 (Check Weekly Report)	Task 4 (Edit activities)
1	√	√	√	√
2	√	√	√	√
3	√	√	√	√
4	√	√	√	√
5	√	√	√	√
6	√	√	√	√
7	√	√	√	√
8	√	√	√	√
9	√	√	√	√
10	√	√	√	√
11	√	√	√	√
12	√	√	√	√
13	√	√	√	√
14	√	√	√	√
Success	**14**	**14**	**14**	**14**
Completion Rate	**100%**	**100%**	**100%**	**100%**

TABLE 3. Task Completion Rates

4. Discussion

Irrespective of the overall assessment of the Daylio app, a close scrutiny revealed some issues with the app. Average time for all participants to complete task 1 was 0.94 minute. Six participants were successful in completing task 1 (checking the status of their mood) below one minute. Other participants had problem when checking their status due to their being unfamiliar with the application. Participant 13 had hard time checking her status (2 minutes) because she was unable to understand some terms in English. She was more familiar with Bahasa Melayu. However, she tried to identify the letters of 'check status' on the interface and was successful at last. Age factor seem to affect the result for this task as well. For instance, participant 10 and 11 which were above 40 years old used more than 1.5 minutes to achieve task 1. Elderly participants took longer time to get familiarized with the interface and to find the correct button to check status. The participants who are familiar with mobile application took shorter duration (Participant 2, 4, 6, 9). The average duration for all participants to complete task 2 (add new entries) was 0.64. When carrying out this task, participants 1, 2, 4, 5, 7, 10, and 11 found it difficult to notice the add mood button. They were not familiar with add (+) button at the middle of interface. However, Participant 3, 6, and 8 had no problem identifying the add (+) button used to add new entries. When pressing add (+) button, the interface showed 5 emoticons that represent mood on that day (good, bad and awful). These emoticons helped users from all over the world to understand how to record their mood quickly without looking to the words. Furthermore, the average duration for all participants to complete task 3 (check weekly report) was 0.81 minute. All participants who took longer time (participant 1, 7, 8 and 10) were iOS users. Daylio application on iOS gave the participants problem to identify the menu or button in order to check weekly report. Its located is hidden behind the menu making these participants to attempt pressing related button to find the weekly report. For android users, they had no problem because the check weekly report button is transparent and easy to reach. In addition, the average time taken for all participants to complete task 4 (edit activities) was 0.93 minute. Edit activities button is easy to find and was clearly located at menu section. However, participant 5, 8, 9, 10, 11 and 13 took longer duration to complete the task. The main reason was that there were lots of symbols representing activities to be chosen. The participants may have taken longer time to choose their correct activities and then tried to add activity for record purposes. These too many symbols may lead to participants' confusion and may tend to make them give wrong inputs.

ISSN: 2005-4238 IJAST

5. Recommendations

App designers want to design apps that bring profit to any business venture and also promote themselves as designers of quality products. Thus, it is necessary to consider fixing the observed issues in Daylio mobile app in order to enhance its usability quality. Firstly, the sign-up screen should be fixed. Presently, it is actually unwelcoming and unattractive to users. First impressions matters and could affect experience of users and lead to product abandonment and user disloyalty. The sign-up forms look longer on mobile platform than on desktop. So designers should to keep it short and capture only very essential information. In addition, the designers of the Daylio app will need to work on the In-app Permissions in some scenarios. These permissions can potentially scare people away if they appear invasive and/or irrelevant as users are apprehensive about handing over access to their personal information, location, social media accounts etc. Besides, the designers may also add controls and transparency to sharing options. Users have accepted the sharing of personal information on mobile devices and within apps, but this has also increased concerns about how those information are stored, shared and used. To establish trust and transparency with users, users should be provided explanation on what is shared and with whom it is shared. All settings should not be automatically set to "allow" as this may make new users feel as if their trust is been violated. It is good to remind users about their sharing settings through an in-app message.

This notwithstanding, designers should standardize the Daylio app interface for Android and iOS platforms. Some buttons on the iOS platform are obscure and somehow hidden. This will cause users to encounter hard times finding some features in the app for example, checking weekly report. Also, there are icons or symbols in Daylio app that are tricky and not easily understandable. The designers should keep app and make icon simple and clear. Furthermore, the icons are visually crammed up because of their so many colors. Though colors and graphics are good, however, designers should focus on the core concept of the app. The icons should be kept simple. The jamming of every app feature in a single icon affects their look and feel and the overall impression about the app.

6. Conclusion

This study examined the effectiveness, efficiency and reliability in use of a mood tracker app. The researchers conducted an evaluation to measure the task completion (success) rate, task completion time and task errors of study participants. This is necessary as the above metrics are measures for effectiveness, efficiency, and reliability in use respectively. Overall, the outcomes reveal that the Daylio mobile app is effective, efficient and reliable in use. However, a close scrutiny indicated that there are pressing issues that need to be fixed to further enhance the quality of the app. Fixing these issues will certainly improve the performance of the app, its market penetration, as well as users/customers loyalty and retention.

References

[1] Hussain, A., Mkpojiogu, E.O.C., Nabeel, N., & Alathwari, A. (2019). Users perception of the mobile usability of a global bicycle sharing platform. *International Journal of Interactive Mobile Technologies (IJIM)*, 13 (1), 125-136.

[2] Breuch, L.-A., Mark Z., & Spinuzzi, C. (April, 2001). "Usability Instruction in Technical Communication Programs". Journal of Business and Technical Comm. 15 (2), 223–240. doi:10.1177/105065190101500204. Retrieved 13Aug 2019.

[3] Hussain, A., Mkpojiogu, E.O.C., Ishak, N., Mokhtar, N., & Ani, Z.C. (2019). An interview report on users' perception about the usability performance of a mobile e-government application. *International Journal of Interactive Mobile Technologies (IJIM)* 13 (10), 169-178

[4] Duh, H. B.-L., Tan, G. C. B., & Chen, V. H. (2006). Usability evaluation for mobile device. In *the 8th conference on Human-computer interaction with mobile devices and services* (pp. 181–186). https://doi.org/10.1145/1152215.1152254

[5] Hussain, A., Mkpojiogu, E.O.C., & Suleiman, K. (2019). A mobile usability testing of Achik.biz application. *Journal of Advanced Research in Dynamical and Control Systems (JARDCS)*, 11(08-SI), 1024-1028.

[6] Dumas, J. S., & Redish, J. (1999). *A practical guide to usability testing*. Intellect books.

[7] Hussain, A., Mkpojiogu, E.O.C., Abubukar, H., & Hassan, H.M. (2019). A mobile usability test assessment of an online shopping application. *Journal of Computational and Theoretical Nanoscience.*16(5-6), 2511-2516.

[8] Dray, Susan; Siegel, David (March 2004). "Remote possibilities?: international usability testing at a distance". Interactions. 11 (2): 10–17. doi:10.1145/971258.971264.

[9] Hussain, A., Mkpojiogu, E.O.C., Faromiki, J.O., Abdusalam, E.M.A., & Shamala, P. (2019). Determining the comfort of use of a mobile navigation app. *Journal of Advanced Research in Dynamical and Control Systems (JARDCS)*, 11(08-SI), 1144-1146..

[10] Gafni, R. (2009). Usability Issues in Mobile-Wireless Information Systems. In *Proceedings of the 2009 InSITE Conference* (Vol. 6, pp. 1–15). https://doi.org/10.28945/3383

[11] Hussain, A., Mkpojiogu, E.O.C., & Suleiman, K. (2019). A heuristic evaluation of Achik.biz mobile shopping app. *International Journal of Recent Technology and Engineering (IJRTE),* 8(2S2), 123-126.

[12] Gough, D., & Phillips, H. (2003). "Remote Online Usability Testing: Why, How, and When to Use It". Retrived on Aug 18, 2019.

[13] Kaikkonen, A., Kekäläinen, A., Cankar, M., Kallio, T., & Kankainen, A. (2009). Evaluation of user interface design and input methods for applications on mobile touch screen devices. *Lecture Notes in Computer Science*, 243–246. https://doi.org/10.1007/978-3-642-03655-2_30

[14] Hussain, A., Mkpojiogu, E.O.C., Ishak, N., & Mokhtar, N. (2019). A study on the perceived mobile experience of MyEg users. *International Journal of Interactive Mobile Technologies (IJIM),* 13 (1), 4-23.

[15] Jerz, D.G. (2000). "Usability Testing: What Is It?". Jerz's Literacy Weblog. Retrieved Aug 16, 2019.

[16] Nadzir, M., Hussain, A., Mkpojiogu, E.O.C., Faromiki, J.O., & Abdusalam, E.M.A. (2019). The effectiveness and efficiency of a GPS route and voice navigation app, *International Journal of Innovative Technology and Exploring Engineering,* 8 (8S), 425-427.

[17] Nielsen, J. (1994). Usability Engineering, Academic Press Inc, p 165

[18] Nadzir, M., Hussain, A., Mkpojiogu, E.O.C., Adetunmbi, S., & Hassooni, M.N. (2019). Assessing the comfort in use of a flipboard mobile news magazine app. *International Journal of Innovative Technology and Exploring Engineering*, 8(8S), 428-430.

[19] Paternò, F., Russino, A., & Santoro, C. (2007). Remote evaluation of mobile applications. In *International Workshop on Task Models and Diagrams for User Interface Design* (pp. 155–169).

[20] Hussain, A., Mkpojiogu, E.O.C., Tam, K.L., Hilaluddin, K., & Jamaludin, S. (2019). A measure of the effectiveness and efficiency of an m-cinema app. *International Journal of Recent Technology and Engineering (IJRTE),* 8(2S2), 127-130.

[21] Ryu, Y. S. (2005). Development of Usability Questionnaires, (July), 1–206.

[22] Technologie, É. De, & Université, S. (2012). THE STATE OF THE ART OF MOBILE APPLICATION USABILITY EVALUATION Fatih Nayebi , Jean-Marc Desharnais , Alain Abran. In *In 2012 25th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)* (pp. 1–4).

[23] Hussain, A., Mkpojiogu, E.O.C., Jamaludin, S., Hilaluddin, K., & Nathan, S. (2019). Users' perception of their satisfaction and experience on a mobile cinema application, *International Journal of Recent Technology and Engineering (IJRTE),* 8(2S2), 131-134.

[24] Andreasen, M.S., Nielsen, H.V., Schrøder, S.O., & Stage, J. (2007). What happened to remote usability testing?. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '07. p. 1405. doi:10.1145/ 1240624. 1240838. ISBN 9781595935939.

[25] Virzi, R. A. (1992). "Refining the Test Phase of Usability Evaluation: How Many Subjects is Enough?". Human Factors. 34 (4): 457–468. doi:10.1177/001872089203400407.

[26] Hussain, A., Mkpojiogu, E.O.C., Karmal, F.M., & Lateef, H.M. (2019). Ascertaining the UX of the word mania mobile app for children using fun toolkit v3. *International Journal of Recent Technology and Engineering (IJRTE)*, 8(2S2), 202-205.

[27] Zhang, D., Adipat, B., & Zhang, D. (2005). Challenges , Methodologies , and Issues in the Usability Testing of Mobile Applications Challenges , Methodologies , and Issues in the Usability Testing of Mobile Applications. *International Journal of Human-Computer Interaction*, *18*(3), 293–308. https://doi.org/10.1207/s15327590ijhc1803

[28] Hussain, A., Mkpojiogu, E.O.C., Adetunmbi, S., Hassooni, M.N., Basri, S. (2019). An evaluation of the effectiveness in use and efficiency in use of a mobile news aggregator magazine application. *International Journal of Innovative Technology and Exploring Engineering*, 8(8S), 527-531.

[29] Balagtas-Fernandez, F., & Hussmann, H. (2009). A methodology and framework to simplify usability analysis of mobile applications. *ASE2009 - 24th IEEE/ACM International Conference on Automated Software Engineering*, 520–524. https://doi.org/10.1109/ASE.2009.12

[30] Jauhar, J., Ghani, A.B.A., Joarder, M.H.R., Subhan, M., Islam, R. (2015). Brain Drain To Singapore: A Conceptual Framework Of Malaysians' Diaspora. Social Sciences (Pakistan), 10 (6), pp. 702-711.