The final publication is available at Springer via https://doi.org/10.1007/978-3-319-60492-3_27

A Hybrid Evaluation Approach and Guidance for mHealth Education Applications

Tareq Aljaber and Neil Gordon

School of Engineering and Computer Science, University of Hull, Hull, HU6 7RX, UK
{T.ALJABER@2013., n.a.gordon@}hull.ac.uk

Abstract. Mobile health education applications (MHEAs) are used to support different users. However, although these applications are increasing in number, there is no effective evaluation framework to measure their usability and thus save effort and time for their many user groups. This paper outlines a useful framework for evaluating MHEAs, together with particular evaluation metrics: an efficient hybrid of selected heuristic evaluation (HE) and usability evaluation (UE) factors to enable the determination of the usefulness and usability of MHEAs. We also propose a guidance tool to help stakeholders choose the most suitable MHEA. The outcome of this framework is envisioned as meeting the requirements of different users, in addition to enhancing the development of MHEAs using software engineering approaches by creating new and more effective evaluation techniques. Finally, we present qualitative and quantitative results for the framework when used with MHEAs.

Keywords: Evaluation Framework · Usability Evaluation · Heuristic Evaluation · Metrics · Stakeholders.

1 Introduction

Enhancements in mobile technology have enabled the development of a wide range of applications that can be used in many aspects of people's lives [1]. An example of these mobile applications is mobile health education applications (MHEAs). MHEAs are widely utilized by a range of different stakeholders. Moreover, these different stakeholders vary in the type of knowledge and background they possess: some are specialists and some are not (mixed users). These different stakeholders need to be able to use the software easily, which is referred to as considering the usability and having an effective framework for the applications. MHEAs are utilized to develop the knowledge of health professionals and patients in order to improve health [2], [3]. Patients need to expand their own health education to ensure consistent development of their health, with support for their well-being coming with MHEAs. Learning and training around health difficulties are progressively substantial issues, mainly as more people are living longer; long life enlarges the scale and complexity of the care of health and well-being [5]. Several areas are covered by health education, such as learning how to manage life in an acceptable way, how to oversee health requirements without difficulty and how to obtain appropriate treatment [6]. In addition to requirements from

the patient side, health professionals can enquire about updating their own medical knowledge and other needs.

The small size of a mobile phone can, however, cause various problems, such as battery life and interface restrictions, which reduce the user experience as well as the quality of the service [8]. Contemporary research by the National Health Service (NHS) has highlighted that mobile health (mHealth) apps can have risks corresponding to the mechanical failure of supplementary medical devices, such as faulty design, and user error, among other safety issues [9]. These faults occur as an outcome of the number of MHEA software developers who do not apply an appropriate framework/model in the evaluation of the usability of their apps to certify that they meet certain requirements, such as enabling health education for patients and health professionals. This emphasizes the lack of an effective framework to evaluate these apps to safeguard the meeting of the requirements of the different stakeholders [2], [3], [4]. The authors of recent work [10], [11] have shown the significance of using heuristic evaluation. Moreover, in this research, we are not looking to the effectiveness of MHEAs in improving health; we are considering their usability as an interface.

2 Why Constructing a Framework is Important

In this second section of the paper, we now consider why there is a need to construct a framework for MHEAs. First, the focus of numerous researchers on creating usability evaluation (UE) metrics in recent decades was not specifically on MHEAs, it was on general systems and principally on web applications [20, 21]. Second, Coursaris and Kim underline in their research [17] the lack of empirical research into the importance of user characteristics as a central point and the impact of the environment on mobile usability. However, they indicate in their paper that the usability possibilities considered in mobile studies have had no subjective research [17]. This highlights the lack of a structured approach to measuring, enhancing and ranking the usability of mobile applications in general and MHEAs in particular. Third, usability is a progressively significant area of development for mobile phones, as it is essential to avoid certain problems [22]. Fourth, and most importantly, according to Coursaris and Kim [17], within the context of this type of mobile computing environment, there is no appropriate UE framework. This demonstrates the necessity of our research, which contains the construction of an evaluation framework for MHEAs. Moreover, as sales have risen sharply (by 10% during 2012/2013 [16]), this has seen the development and usage of mHealth applications growing markedly since 2013: over 36% of all mHealth applications in 2014 were released the previous year [15]. According to Dubey et al. [14], usability inspections benefit from certain types of methods, such as surveys, logging, and interviews. Lastly, the number of mHealth apps is rising at a significant rate: more than doubling in the year between 2015 and 2016 by 57% to 259,000 apps [7].

2.1 The Importance of Usability for Mobile Health Education Applications

Mobile applications are a significant development area in current computing [12]. According to Hernandez-Munoz and Woolley [13], mobile phones are fundamental

devices for the general population, which indicates that many people depend on them for a wide range of aspects of their daily lives, such as transportation and health education. Recent investigation demonstrates that one in every five people in the world uses a smartphone [18]. Current estimates are that in 2014 about 497 million mobile phones were in use and that smart devices characterized 26% of these [19]. This suggests the increasing usage of mobile phones, reflecting a wide growth in the potential for having extra applications utilized. This emphasizes the significance of being able to rank and distinguish the usefulness of apps. One of these vast areas of applications is MHEAs, which are utilized by various stakeholders from different backgrounds for several aspects of their lives under the umbrella of health education. This emphasizes the need to identify what the different stakeholders require, in order to be able to measure MHEAs from the viewpoint of usability.

3 The Methodology of Selecting Usability Metrics for the Framework

Heuristic evaluation (HE) and UE have been applied in many respects [2], [10]. In the early stage of our research, we depended on the literature in selecting usability metrics on the method of general sets have been hybrid selected from HE and UE [2]. Moreover, we proposed a specific set depending on the literature and then made a hybrid selection of metrics taken from HE and UE metrics, proposed and filtered by conducting 15 selected interviews with health professionals (HP), software developers (SD) and patients (P).

With the original selection of metrics, an analysis of data from the 15 interviews (with HPs, SDs and Ps) demonstrated the necessity of modifying some of the metrics originally selected because they did not fully match the user requirements identified during the interviews. Some of the metrics were not necessary to the requirements, although some of the metrics were required to be complementary to our proposed hybrid selection.

3.1 Hybrid Metrics

We categorized and systematized our hybrid selection of HE and UE metrics as follows: A) Memorability, B) Features, C) Attractiveness, D) Simplicity (containing learnability), and E) Accuracy. These metrics are clarified in Figure 1 [3].

Through the performance of the 15 unstructured interviews, we obtained an improved hybrid of five selected metrics: (A) Memorability, (B) Features, (C) Attractiveness, (D) Simplicity (containing learnability), and (E) Accuracy. Obtaining an improved hybrid of selected metrics supported us in mapping actual questionnaire questions to specific metrics, the intention being to allow us to assess usability issues for MHEAs accurately. Moreover, the interview questions were designed for two outcomes: the ranking of existing MHEAs and as an aid to designing a prototype.

4 Hybrid Evaluation Framework

The framework passed through three design stages as the research progressed: the first design is illustrated in Figure 1 [2]; the second design is presented in Figure 2 [3]; and the third design is shown in Figure 1 [4]. This paper focuses on the third design.

The third design of the framework includes the first and second designs, modifying the last part of the framework as a tool to guide the selection of MHEAs based on usability criteria. This design imitates the first two experiments we carried out in our framework: first, unstructured interviews were conducted with different stakeholders (HPs, Ps and SDs), which directed us to a hybrid of metrics selected to meet the stakeholders' requirements; second, a questionnaire was distributed based on mapping between the results of the unstructured interviews and the questionnaire topics.

However, for this design, the purposes identified from the questionnaire were amended slightly to ranking existing MHEAs, as well as the development of a tool to assist different stakeholders choose between various MHEAs depending on their requirements. The selection tool was based on data from the analysis of the feedback to the questionnaire. The design is shown in Figure 1.

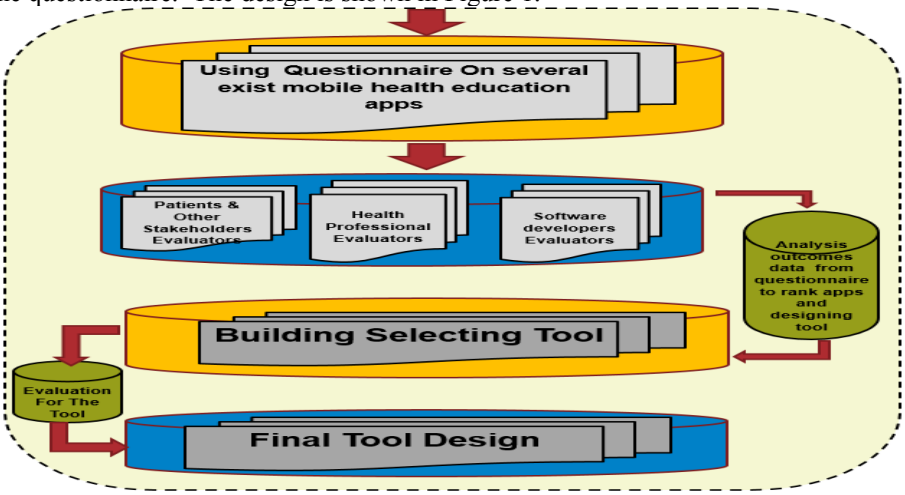

Fig. 1. Evaluation framework.

5 Hypothesis

In this paper, we address two key questions: first, is it possible to construct an effective evaluation framework by developing an efficient hybrid of HE and UE metrics in order to measure the usability of MHEAs? Second, is it possible to use this evaluation framework to identify current issues and effectively improve the design of MHEAs?

The hypothesis that we can develop from the above is whether we can distinguish between MHEA packages using the framework displayed in Figure 1:

H1 (the hypothesis) is that there is a classifiable difference between some MHEA packages.

H0 = null hypothesis = there is no difference between some MHEA packages.

6 Qualitative Comparison Results for the Framework

Following the gathering of data from 15 unstructured interviews (five with health professionals, five with patients and five with software developers), we analyzed the responses from all three sets of interviewees to detect their most important requirements when using MHEAs. Our results are shown in Figure 2 as gathered from all of the stakeholder categories, distributed by HP, P and SD and detailed below.

The most important metric to be measured in meeting the requirements of health professionals, patients and software developers is (D) Simplicity, showing the highest chart bar. In second place is (B) Features, showing the second-highest chart bar. Third is (E) Accuracy, showing the third-highest chart bar. In fourth place is (A) Memorability, showing the fourth-highest chart bar. In last place is (C) Attractiveness, showing the lowest chart bar. These results assist us in stating that the most important metrics for health professionals, patients and software developers when using MHEAs are (D) Simplicity and (B) Features.

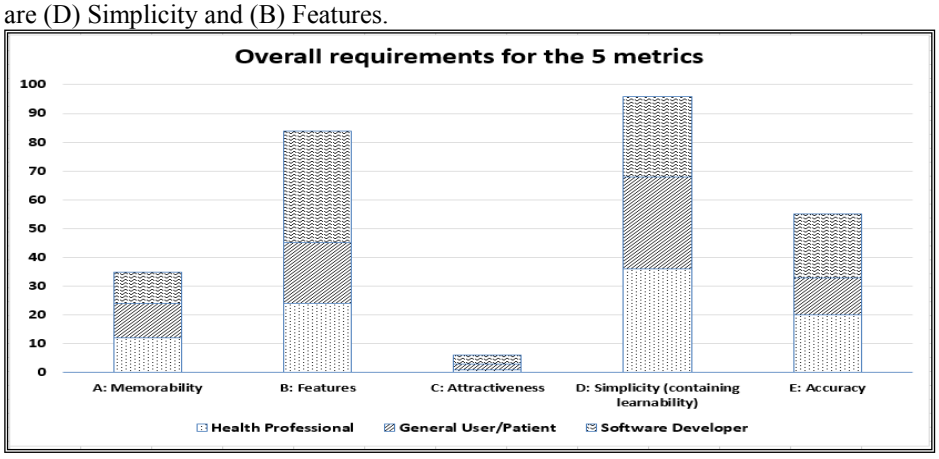

Fig. 2. Interview results measuring the requirements of patients, health professionals and software developers for the five metrics considered

7 Quantitative Comparison Results for the Framework

Following data collected from a questionnaire completed by 81 participants distributed into equal sample sizes for each of the three categories (27 health professionals, 27 patients and 27 software developers), we obtained the outcomes from the analysis of the overall feedback, identifying that different stakeholders preferred different metrics within the same MHEA. The radar plot in Figure 3 demonstrates that health professionals – of all the stakeholder categories – most preferred using Medscape for its (A) Memorability, (B) Features and (D) Simplicity. Patients/general public were in second place, most preferring to use Medscape for the (C) Attractiveness metric. Software developers were in third place for most preferring to use Medscape for the (E) Accuracy metric. Furthermore, from the T-test we carried out between HPs and SDs, we obtained a P-value of 0.042, which is less than 0.05. Thus there is a statistically significant difference between HPs and SDs with regard to what metric they preferred.

This suggests that we can differentiate between the different stakeholders in terms of what metrics they prefer.

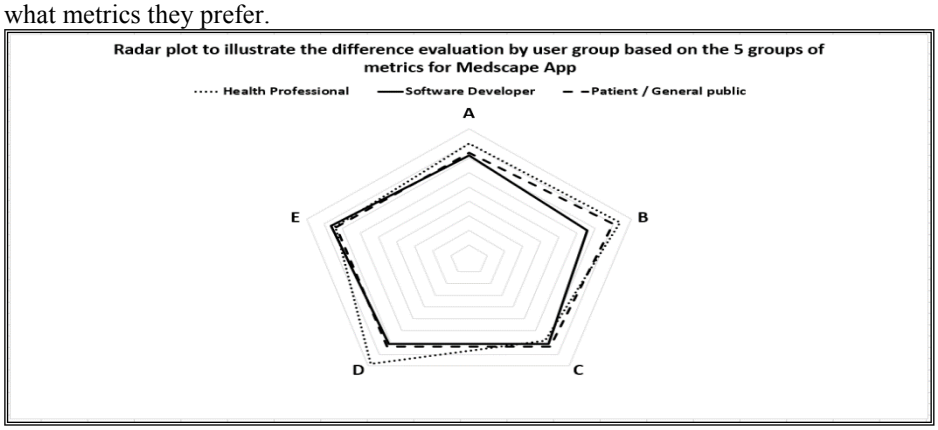

Fig. 3. Medscape app ranking the comparison of the five metrics considered by HP, SD and P/General public

The whisker chart in Figure 4 illustrates some of the results from our work in applying the framework. Observing the patterns in the whisker chart, we can identify visual differences between the three applications, providing evidence that the framework can differentiate between these three types of MHEA. Moreover, focusing on the first MHEA - Epocrates - on the A) Memorability metric, we can see that the average is 3.5, the median is 3, the lower quartile is 2 and the upper quartile is 5. Furthermore, by comparing the A) Memorability metric pattern from Epocrates with the other MHEA A) Memorability patterns, we can identify the differences between the average, median, lower quartile and upper quartiles. Finally, from the T-test between the A) Memorability metric from Epocrates with the A) Memorability metric from Medscape, we obtained a P-value of 0.045, which indicates that there is a significant difference between the two MHEAs within the A) Memorability metric. This shows that we can differentiate between these MHEA packages.

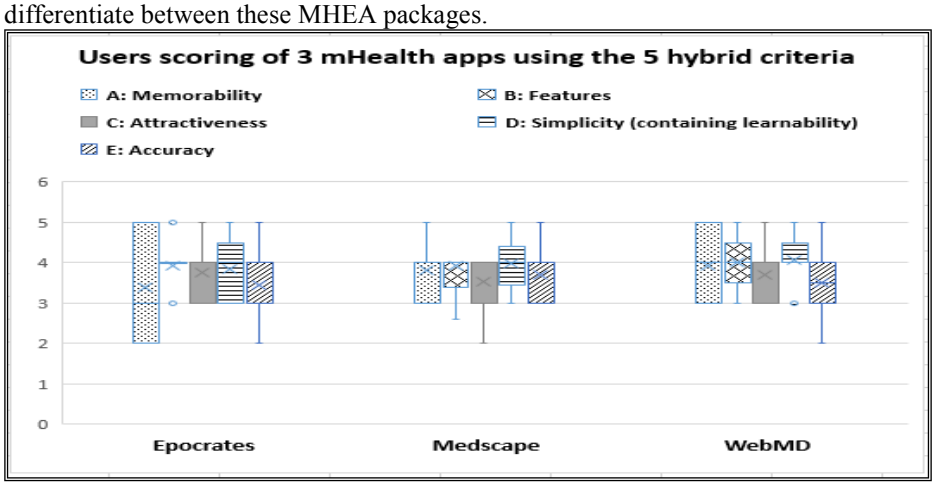

Fig. 4. Comparison of all users' scoring of three MHEAs for the five metrics considered

8 Conclusion and Contribution

Different stakeholders utilize MHEAs for various critical purposes. If an MHEA does not offer users the opportunity to save time and effort, or simplify their understanding of health matters, it will not be utilized effectively. This is part of the broader goal of satisfying the user requirements of patients, health professionals and their different partners. This paper has demonstrated how HE and UE metrics can offer an effective technique to solve the above-stated problem.

Moreover, this paper has explored the ability to improve MHEAs through evaluating them in the initial stages of software design. In this case, our hybrid of five selected usability metrics - Memorability, Features, Attractiveness, Simplicity and Accuracy - can be enhanced to expand the probability of these apps being successful and useful. This type of framework can be employed as an evaluation tool for several types of mobile app. The focus in this paper is on highlighting the structure of the evaluation framework for MHEAs by depending on a hybrid selection of metrics (HE and UE), combining five metrics selected from a larger set of heuristic and usability evaluation factors, and examining them based on interviews with health professionals, patients, and software developers. The five metrics parallel specific aspects of usability acknowledged through a requirements analysis of typical users of mobile health apps.

The five metrics were broken down into 24 specific questionnaire questions. The intention of this project is to construct a set of tools for the evaluation of MHEAs to enable them to be ranked, and thus assist different stakeholders (such as patients and health professionals) to choose apps that are most suitable for them. Moreover, the evaluation framework proposed can offer guidance on system specifications as well as different user requirements for the software developers of mobile health apps. Finally, we concluded some quantitative and qualitative results in this paper by highlighting the ability to distinguish between MHEAs from the usability side using our framework, which will help us in further research steps to develop a prototype for the guidance tool.

References

1. Harrison, R., Flood, D., Duce, D.: Usability of mobile applications: literature review and rationale for a new usability model. J. Interact. Sci. 1, 1–16 (2013).
2. Aljaber, T., Gordon, N., Kambhampati, C., Brayshaw, M.: An evaluation framework for mobile health education software. In: 2015 Science and Information Conference (SAI). pp. 786–790. IEEE, London (2015). doi: 10.1109/SAI.2015.7237233
3. Aljaber, T., Gordon, N.A.: Evaluation of mobile health education applications for health professionals and patients. In: 8th International conference on e-Health (EH 2016). pp. 107–114. IDIAS, Funchal (2016).
4. Aljaber, T., Gordon, N.: A Guidance and Evaluation Approach for mHealth Education Applications. (Accepted). In: 2017 human computer interaction international 2017 Conference (HCII 2017). Springer, Vancouver, Canada (2017).
5. Shareef, A.F.: Special issue on 'Innovation in Distance Learning Technologies in Developing Countries'. J. Learn. Technol. Newsl. 8, 1–27 (2006).
6. Glanz, K., Rimer, B.K., Viswanath. K.: Health Behavior and Health Education: Theory, Research, and Practice. Jossey-Bass Wiley, San Francisco (2008).

7. research2guidance 2016: mHealth App Developer Economics 2016. Technical report, Berlin (2016).
8. Silva, B.M.C., Rodrigues, J.J.P.C., Lopes, I.M.C., Machado, T.M.F., Zhou, L.: A Novel Cooperation Strategy for Mobile Health Applications. IEEE J. Sel. Areas Commun. 31, 28–36 (2013).
9. Applications, H.: App Development: An NHS Guide for Developing Mobile NHS Innovations South East. Technical report, (2014).
10. Gordon, N., Brayshaw, M., Aljaber, T.: Heuristic Evaluation for Serious Immersive Games and M-instruction. In: Learning and Collaboration Technologies. pp. 310–319. Springer International Publishing, Toronto (2016). doi: 10.1007/978-3-319-39483-1_29
11. Brayshaw, M., Gordon, N., Nganji, J., Wen, L., Butterfield, A.: Investigating Heuristic Evaluation as a Methodology for Evaluating Pedagogical Software: An Analysis Employing Three Case Studies. In: 1st International Conference on Learning and Collaboration Technologies. pp. 25–35. Springer International Publishing, CRETE (2014).
12. Zemliansky, P., St. Amant, K.: Handbook of Research on User Interface Design and Evaluation for Mobile Technology. Information Science Reference, New York (2008).
13. Hernandez-Munoz, L.U., Woolley, S.I.: A user-centered mobile health device to manage life-threatening anaphylactic allergies and provide support in allergic reactions. In: 9th International Conference on Information Technology and Applications in Biomedicine (ITAB). pp. 1–4. IEEE, Larnaca (2009).
14. Dubey, S.K., Rana, A.: Analytical Comparison of Usability Measurement Methods. Int. J. Comput. Appl. 39, 11–18 (2012).
15. research2guidance 2014: mHealth App Developer Economics 2014. Technical report, Berlin (2014).
16. Smith, A.: Smartphone Ownership – 2013 Update. Technical report, Washington (2013).
17. Coursaris, C.K., Kim, D.J.: A Meta-Analytical Review of Empirical Mobile Usability Studies. J. Usability Stud. 6, 117–171 (2011).
18. Heggestuen, J.: One in Every 5 People in the World Own a Smartphone, One in Every 17 Own a Tablet, http://www.businessinsider.com/smartphone-and-tablet-penetration-2013-10?IR=T.
19. CISCO: Cisco Visual Networking Index: Global Mobile Data Traffic Forecast Update 2014–2019 White Paper, https://ec.europa.eu/futurium/en/content/cisco-visual-networking-index-global-mobile-data-traffic-forecast-update-2014-2019-white.
20. Alva, M.E.O., Martínez P., A.B., Cueva L., J.M., Sagástegui Ch., T.H., López P., B.: Comparison of Methods and Existing Tools for the Measurement of Usability in the Web. In: International Conference on Web Engineering (ICWE 2003). pp. 386–389. Springer, Berlin (2003).
21. Ivory, M.Y., Hearst, M.A.: The State of the Art in Automating Usability Evaluation of User Interfaces. J. ACM Comput. Surv. 33, 470–516 (2001).
22. Baharuddin, R., Singh, D., Razali, R.: Usability Dimensions for Mobile Applications-A Review. Res. J. Appl. Sci. Eng. Technol. 5, 2225–2231 (2013).