Automated User Experience (UX) Testing for Mobile Application: Strengths and Limitations

Hayfa. Y. Abuaddous[1], Ashraf Mousa Saleh[1(✉)], Odai Enaizan[2], Fahad Ghabban[3], Anas Bassam Al-Badareen[4]

[1]Faculty of Computer Sciences and Informatics, Amman Arab University, Amman, Jordan
[2]Department of Management Information System, University of Tabuk, Tabuk, Saudi Arabia
[3]College of Computer Science and Engineering, Taibah University, Medina, Saudi Arabia
[4]Department of Software Engineering, Aqaba University of Technology, Aqaba, Jordan
`asaleh@aau.edu.jo`

Abstract—User Experience (UX) evaluation has a significant importance for any interactive application. Mobile device applications have additional limitations to convey good user experiences (UX) due to the usage and features of mobile devices in ubiquitous computing contexts. There are a lot of automated tools available for evaluating, such applications particularly developed for Android and iOS. Automated UX evaluation provides a number of strengths which are able to overcome the limitations of non-automated (moderated) counterparts. Automated UX evaluation is regarded as unexplored field, especially when it is adopted for the UX evaluation of complex mobile application. Despite these, automated UX evaluation tools are unexplored and should only be used in coupling with traditional UX evaluation techniques. Hence, this study aims to discuss the literature in automated UX evaluation and to highlight the UX measurements that merit further investigation. This paper presents the most frequently used automated UX evaluation tools for mobile applications. Therefore, the study analyzes existing automated UX tools and frameworks and identifies which UX attributes and measurements of automated UX evaluation are likely to be of use in future research. Moreover, we discuss the strengths and limitations of automated UX evaluation methods for mobile applications. Finally, the paper provides an overview of automated UX evaluation techniques and tools. It also presents the strengths and limitations for comparing various evaluation methods.

Keywords—user experience (UX), automated UX evaluation, UX evaluation, moderated UX evaluation, mobile applications

1 Introduction

User Experience (UX) evaluation focuses on how users can learn, use and be satisfied with a product or system. Currently, many companies and industries are motivated to invest in designing or improving their UX applications as customer satisfaction and loyalty have shown a direct impact on those companies' sales and profits. In fact, there

are broad strands of literature, which showed that every dollar spent on UX could bring from $2 and up to $100 dollars in return [45]. There are varieties of methods that gather feedbacks from users on the UX of products or systems including websites, software, devices, or applications. UX evaluation of mobile applications is considered to be as an up-and-coming area of research that associated to the human computer inter- action (HCI) [1]. The primary goals of UX evaluation are to identify the effectiveness, efficiency and satisfaction of the users as well as it could also identify issues that may lead to human errors and user frustration [2]. However, the diversity offered in UX evaluation tools and techniques such as interviews, workshops, usability testing (lab and remotely), contextual research, expert reviews, surveys and automated tools.

The UX evaluation of mobile device applications is slightly different from desktop computer applications as mobile applications have unique features such as limited bandwidth, small screen interface and changing mobile context (i.e. location and limited memory). Thus, the methods and guidelines proposed for the evaluation of desktop computer applications may not be directly applicable to mobile device applications [3–5].

Prior literature had shown that the diversity of UX evaluation methods can greatly impact the evaluation process. For example, UX evaluations which aim to ensure the mobile applications are suitable, fit user requirements, and satisfactory for users had generated different results under the same method of evaluation. According to [6], if the same UX evaluation method was conducted by different evaluators, the results would hardly intertwined. This indicates a lack of systematic approach used to get the eval- uation results besides of involving neither some UX attributes nor other used in UX experiments [44]. In addition, the UX evaluation methods in context of laboratory or field study have also shown to influence the evaluation results. Thus, automated UX evaluation can provide a solution since for this issue as it can be applied in both field and laboratory setting. Automated UX evaluation such as analytics can gather mas- sive information and establish the base for better designing and management decisions. Automated UX evaluation can also help in identifying users' behavior through evaluat- ing their quantitative data. Consequently, the automated UX evaluation becomes able to understand user behaviors when they are using the product. The automated tools report has highlighted major issues regarding leaky conversion paths or drop off on product pages [46].

An automated UX evaluation transparently uses log UX metrics during UX eval- uation and inspection from unaware end-users. In addition, the user is not guided by evaluators about who could influence the study of their data by foreclosing interaction possibilities. [7] claimed that automated UX is a trend due to its various advantages, such as saving time and efforts of evaluators, and increase the opportunities to detect error of evaluated UX attributes and metrics. Despite the potential advantages of auto- mated UX evaluation, it is important to notes that automated UX evaluation is a useful complement and has contributed to the traditional UX evaluation techniques in many aspects such as expert review and user research.

Different evaluation techniques can uncover different types of UX attributes from both objective and subjective measurements. For example, users satisfaction are unlikely to be covered by automated methods [7]. Automated tools involves in quantitative data, based on statistically significant amounts of data. The review of UX evaluation of

mobile application researches have shown that proposed methods are adequate for evaluation process. Although many studies have investigated the UX evaluation methods, there are still some gaps regarding empirical evidence. More precisely, the selection of empirical methods for UX evaluation is still missing in the literature. The fact that the user is the center of application design has necessitated that the automated evaluation method to consider different tasks and context-of-use for the evaluation experiment. Moreover, UX attributes and UX metrics have a considerable impact on the automated UX evaluation performance. Thus, this study addresses the new issues that arise in UX evaluation in mobile application context. The current study also sheds the lights on the existing UX evaluation tools for mobile applications which were adopted by different studies; Hence, this paper aims to:

1. Highlights the automated UX evaluation tools and frameworks which support evaluators in carrying out user interaction evaluations in mobile application. Therefore, the study conducts a review of the automated UX evaluation tool in terms of effectiveness, efficiency, learnability, memorability, errors, satisfaction, cognitive load, and simplicity to fill the gap in the literature.
2. Identify the strengths and weaknesses of automated UX evaluation regarding the experiment task, user, and context of use. The current study also focuses on three defined UX factors, namely user, task, and mobile context.

The remaining of this paper is structured as follows: section 2 discusses related work in the field of moderated UX evaluation as traditional method; section 3 discusses related work to automated UX evaluation; section 4 presents the methodology of this research, Results and Discussion are presented in section 5. Finally, section 6 concludes the paper and emphasizes the main ideas then future works in the area of automated UX evaluation are introduced. Figure 1 illustrates the study block diagram for automated UX assessment tool.
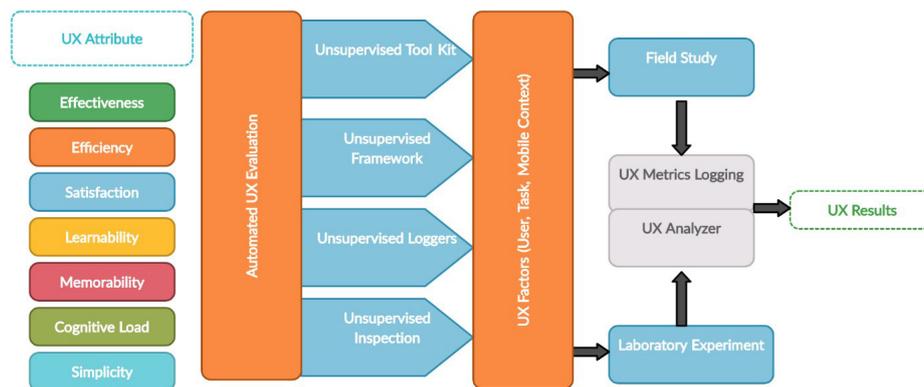

Fig. 1. Automated UX assessment tool block diagram

2   Moderated UX evaluation method

Traditional and moderated UX evaluation methods use heuristics evaluation, cognitive walkthrough (CW), field studies and laboratory studies. The traditional UX evaluation methods which guided by evaluators are applicable to both desktop computer and mobile device applications. However, these evaluation methods are struggling in term of achieving relevant and statistical significant results especially for mobile device applications. This may refer to the fact that emulating real-world uses in a laboratory-based evaluation is only feasible when the user context is precisely defined. As a result, the lack of various contexts and physical constraints and the difficulty of generalizing to a rapidly evolving and probably strongly varying user context [8].

Heuristic evaluation is an informal UX method commonly used for finding major UX problems in a user interface. Therefore, Heuristic evaluation is considered a type of UX inspection method that examines and evaluates the user interface for UX issues [9].

A small set of expert's evaluators examine the interface and evaluates its compliance with recognized UX principles. As a certain level of knowledge and experience is required, it is difficult to find UX experts and employing them is expensive, particularly if they should have domain expertise [10].

UX is not the most applicable method of evaluation when it comes to performance measurements [11, 12]. UX can influence the users behavior as verbalization is required throughout the performance [2]. In this case, moderators must monitor and record the behavior of users, as the user's verbal protocols may often be contrived, biased and misleading which may lead to unreliable results [2].

The CW scrutinizes the simplicity of tools. In other words, CW is able to determine how easy it is for a new or inexperienced user to operate a given tool. The user's knowledge of the domain and task is also considered throughout the evaluation.

In addition, CW evaluation attempts to identify to which extent a user interface supports the novice user in understanding and learning how to achieve tasks through their interfaces [13]. CW requires expert's evaluators and an intensive training of novice users to conduct the evaluation. However, when the task is relatively large or complex, it can be a time consuming.

Analogously, there are a variety of weaknesses affiliated with the CW evaluation. For instant, the fact that evaluation under CW can detect very detailed issues such as how the buttons are marked, whereas, ignoring more generalized issues such as the menu structure and the user interface sequence can raise an intriguing question about the reliability of the evaluation [3, 14].

The studies of [15] and [16] also revealed that field studies require twice the time and effort compared to laboratory studies due to the required preparation of test participants and their supervision. Therefore, CW as a method for evaluating the UX of mobile devices is not recommended.

Conducting the pre-tests or pilots using the traditional UX evaluation is critical process. This assumption was formulated based on the UX evaluators who have to supervise the evaluation method in the field and laboratory studies. Another major weakness of field experiment is that it is require expertise not only in ethnography, but also in the field understudy. Moreover, it is difficult to translate the findings of qualitative field studies into design.

Traditional UX evaluation requires an expert to conduct the evaluation and there is a possibility of having a shortage of such experts [9]. Experts are also costly which make this method neither feasible nor cost effective [17, 47]. These barriers surrounding the UX evaluation method led to the concept of automated and automated UX evaluation.

3   Automated UX evaluation method

Xu et al. [18] highlights that automated UX tools are more efficient and effective in user testing as these tools can save time and cost. They also suggest that automated UX tools can maintain high level of consistency and quality in application designing. An UX evaluation transparently uses log UX metrics during UX evaluation and inspection from unaware end-users without being guided by evaluators who could influence the study by foreclosing interaction possibilities [19].

Dingli & Mifsud [20] listed a number of benefits of using an automated UX evaluation compared to the traditional UX evaluations, they are summarized as follows:

1. Cost reduction: The UX evaluation processes can be expensive since it requires laboratory preparation, testing environment, and UX evaluators and expert's supervision. Time spent on designing tests and in the case when expert evaluators employed to conduct UX testing causes additional costs incurred. Early and frequent testing using automated processes in the development stage can save cost. The developers have the chance of detecting issues in timely bases for corrections. Automate Methods can also capture, analysis, or critique activities leading to time and cost saving. Finally, automated remote UX evaluation does not require a direct supervision from the UX evaluator.
2. Competitive edge: Traditional UX evaluation methods can be time consuming. If designers a reattempting to introduce a new product to the market within a limited period of time while undergoing all its necessary testing, automated tests would be most suitable as they are much quicker to undergo than traditional UX evaluation methods, allowing application designers to detect issues and fix it in a more time efficient manner.
3. Regression testing: regression testing has been gaining popularity recently and became important for developers. Regression testing ensures that no new errors or flaws are caused to the already existing system and its functions. An automated test that yields consistent results can support in ensuring that an application improved over time.
4. Comparative testing: designers usually choose between applications for deployment. Automated testing helps developers to complete an accurate, side by side comparison of two or more applications.
5. Meeting demands: Automated tools are more capable than human UX evaluators in meeting large market demand of application, whereas, designers and developers are more suitable for small projects. The main benefits of using automated tests are efficiency and repeatability. Both benefits are directly related in reducing time and cost. However, the major concern about automated tests is their accuracy.

Automated UX evaluation tools perform data collection during experiment while the UX evaluator captures respondents' facial expressions and other cues. Direct UX data analysis is another feature provided by automated UX that enables the evaluator to draw clear insight into the mobile application.

In this context, automated UX evaluation can be both user-report oriented or automated usage tracking based on UX attributes and metrics determined by UX evaluator. Several UX attributes and metrics can be further added into automated evaluation due to the flexibility of objective and subjective measurements.

However, the existing automated UX evaluation tools and frameworks for mobile application have several limitations such as lack of UX metrics to guide the evaluation of UX attributes and not including user, task, and context as factors during the UX evaluation [40, 41]. The evaluation methods based on observation in a controlled environment without analyzing the effects of unexpected interruptions cannot explicitly measure efficiency, user errors or user satisfaction while using the mobile application.

A UX evaluation requires selecting appropriate UX attributes, factors and adopting suitable evaluation method to obtain relevant UX information. The review of UX evaluation has showed that the existing method has its limitations such as being limited to specific application, not conducting validation or validation and neglecting UX attributes and factors. This shows that various UX attributes and factors can influence the automated UX evaluation method. In this regards, Table 1 presents the main previous studies which tackle this issue.

Some automated tools and frameworks focus only partly on descriptive UX metrics. Thus, it cannot explicitly provide data to identify general UX attributes as identified by [12, 21, 22].

Existing automated UX evaluation tools and frameworks are limited to capture the metrics of user interaction in the absence of clear classification of these metrics in term of UX attributes that defined in UX evaluation models [17, 19, 23–25]. Moreover, these frameworks do not consider the UX attributes as a part of proposed automated frameworks. Consequently, these frameworks are able to track basic user statistics without considering UX factors such as user, task and context during UX evaluation. Moreover, automated UX evaluation cannot explicitly provide data to identify UX metrics such as efficiency, user errors or user satisfaction.

In summary, both researchers seem to recognize that poor UX attributes identification of inappropriate mobile application UX evaluation are expected to lead to a poor application [26, 27, 48]. The consequences of poor mobile application design include poor-quality application, user dissatisfaction and mobile application failure. Moreover, Kiili. [28] found that poor UX of mobile device applications tend to hinder using mobile application goal, again supporting the importance of UX to the success of mobile device applications.

4 Research methodology

The current study included 45 papers related to the UX filed. However, the author sieved out the most related 19 papers that are directly linked to UX automated tools. Each paper was fully reviewed to extract the evaluation method, UX attributes and metric and UX factor. A summary of those studies is provided in Table 1.

Table 1. Automated UX evaluation methods and attributes

Author(s)	Method	Evaluation Method	UX Attributes	UX Factor
Zduniak [30]	Field survey	Automated evaluation tool	–	Context
Balagtas et al. [17]	–	Automated logging framework	Learnability, efficiency, errors, satisfaction, comprehensibility	User Task
Lettner et al. [31]	Field survey	Automated logging of user interaction	Efficiency error rate	User
Artur et al. [24]	Field survey	Automated model	Effectiveness, efficiency and satisfaction	User Task
Bader and Pagano [25]	Field survey	Automated method	Retention time, visitation sequence and visitation history	User
Kluth et al. [19]	Laboratory experiment	Automated capture framework	Users' misses Number of accidental touches	User
Ma et al. [23]	Laboratory experiment	Automated toolkit	Number of backtracks Number of user mistakes Number of correct flows	User Task
Xu et al. [18]	Expert evaluation	UX inspection framework	Simplicity Naturalness Consistency Forgiveness and feedback Effective use of language Efficient interactions Effective information presentation Preservation of context Minimum cognitive overload	Context
Dixit and Padmadas [32]	Laboratory experiment	UX help system	Learnability Navigability	User Task
Scheller and Kühn [33]	Laboratory experiment for developers	API UX concepts framework	Interface complexity for High and low level concepts Learning effects	User
Feijó et al. [10]	Field survey Laboratory experiment	Automated UX testing (Emotion Logger)	Application page User events e.g. Tap emotions levels Emotional events	User Task
Harms [34]	Field survey	Automated approach for UX evaluation	UX of GUI Text format Misleading clicks	User Task
Ferre et al. [35]	Laboratory Experiment	Automatic logging	Low level user action	Task User
Moran et al. [36]	Laboratory experiment	Automated UX tool (GVT)	Efficiency Accuracy	Context Task
Mathur et al. [37]	Laboratory experiment	Automated code analysis framework	Validation Verification	Context Task

(Continued)

Table 1. Automated UX evaluation methods and attributes *(Continued)*

Author(s)	Method	Evaluation Method	UX Attributes	UX Factor
Souza and Canedo [6]	Laboratory experiment Literature review	Automated UX questionnaire	Subjective satisfaction	User Task Context
Ntoa et al. [7]	Laboratory experiment	Automated UX observer tool	Task success Number of help requests User errors	User Task
Sodhar et al. [42]	–	Automated Tool Webpage Analyzer	Accessibility Interoperability	Context
Saleh et al. [43]	Field survey	Automated UX Tool (Interaction logging too)	Efficiency Effectiveness Satisfaction	User Task Context

5   Results and discussion

The analysis of the studies listed in Table 1 shows that the majority of the UX studies conducted in the controlled lab, which is insufficient to evaluate mobile applications in mobility context. Moreover, empirical methods should take into consideration the mobile context of use due to the highly dynamic context of mobile applications, laboratory and field UX experiments encompasses different challenges and can uncover different UX issues [29].

This study has highlighted that even though a number of automated tools and frameworks have been proposed over years for use as automated UX evaluation for mobile application, there is still a lack of adequate and versatile UX attributes that could be employed by evaluators of similar context. It is clear observed efficiency, effectiveness, learnability, memorability, errors, cognitive load and satisfaction corresponds to PACMAD model that those are most investigated attributes in identified UX approaches [21]. For mobile applications, it is incorrect to assume that a UX attribute is less or more relevant to evaluate than another attribute just for being more addressed.

Main observation from this review is that there is an essential need to analyze the perceived importance of discussed UX attribute from user's perception to map UX metrics to their attributes, which can proof importance of these attributes and related metrics for UX evaluation. Moreover, relationships between UX attribute be considered when assigning these attributes for UX evaluations. Results only describe state of art on use of these attributes in UX evaluation.

UX evaluator can analyze data collected at communication and task level from a physical and a temporal distance. This analysis is then feedback to development process and can be used to improve design. In case of automated usage tracking, behavioral usage data are automatically collected by system when user interacts with it. The subsequent data analysis can be either manual or automatic [23].

According to Arthur et al. [24] distinguished two alternatives, possibly simultaneously approaches in term of automated UX evaluation which are a) user-reporting techniques for collecting subjective feedback and b) usage tracking techniques for click stream data behavioral analysis. Results are automatically reported and contain information about aggregate verbatim answers, completion time, abandon frequency, navigation paths and satisfaction indicators. These are most common automated UX evaluation metrics.

Currently, there is a clear cut difference between the responsibilities of developers and UX evaluation specialists. It requires additional management techniques to coordinate the activities of both parties effectively. According to Bader & Pagano [25] and Kluth et al. [19], UX evaluation specialists are replaced by automation tool while test participants are exchanged with consumers of launched application.

UX specialist can sometimes lack the experience or knowledge in the application domain being tested, affecting the validity of their opinions and findings. Furthermore, developers may sometimes base design decisions on their own understanding of the UX evaluation results if they do not play a frequently active role. This means that developers do not gain the knowledge that he/she requires and is less likely to consider UX as an integrated part development. In this case, automation of UX evaluation and testing would be the most accurate solution, in the same manner that functional testing has been automated [38].

Automated UX activities are performed without the direct supervision of the user interactions [6, 7]; in other words, no potential human interference effects on user interaction. Moreover, UX evaluation tools perform data collection during experiment while the UX evaluator captures respondents' facial expressions and other cues. Direct UX data analysis is another feature provided by automated UX that enables the evaluator to draw clear insight about mobile application. The majority of automated UX evaluation tools discussed earlier consider the objective measurement while the subjective measurements such as user satisfaction are rarely covered in automated UX [7].

In the light of the above, available automated UX evaluation tools consider to be a useful complement and addition to traditional UX evaluation techniques. This paper discusses the strengths and weaknesses of available automated-automated UX evaluation in context of mobile application as shown in Table 2. The table shows that the majority of the studies did not adopt standard UX model, such as ISO [39], Nielsen [12] and PACMAD model [21].

Table 2. Automated UX evaluation strengths and weaknesses

Author(s)	Strengths	Weaknesses
Zduniak [30]	– Reusable tool components – Generic for several UX test cases	– Did not involve UX attributes based on UX model. – Recording user inputs – Did not directly relate to UX
Balagtas et al. [17]	– The proposed framework oriented to developer and UX evaluator during mobile application development – Focuses on technical side of the UX evaluation and the collected metrics extracted and formatted to analysis phase	– No evaluation conducted – No validation performed
Lettner et al. [31]	– Tracing any user interaction during mobile application life cycle	– Did not consider the mobile-context as UX factor. – No predefined task was assigned. – Limited to efficiency and error attributes.
Artur et al. [24]	– Hybrid model measuring quantitative, subjective, and context data	– Did not considers sufficient UX attributes suits mobile device application context. – Did not consider the mobile-context as UX factor.
Bader and Pagano [25]	– Conducted the experiment with real mobile application – Proposed lightweight model for automated UX evaluation can be integrated into different mobile applications	– Did not considers the UX attributes suits mobile device application context. – Did not consider the mobile-context as UX factor.
Kluth et al. [19]	– The proposed framework is able to capture and analysed the detected problems	– Limited to identify user interaction problems but it does not consider UX attribute into consideration. – Did not consider task and mobile-context as UX factor
Ma et al. [23]	– The proposed toolkit is able to capture a detailed user interaction supported with quantitative analysis	– Limited to Identify UX problems – Did not determine UX attributes and related metrics in capturing phase. – Limited to efficiency attribute.
Xu et al. [18]	– The UX inspection framework is able to review the guidelines of mobile UX. – Measure the operationalized metrics by translating the UX guidelines	– Did not suggest a universal threshold for any UX metric. – The metrics such as "number of steps of a task" needs to be further studied to see whether there is a good threshold value in terms of UX of mobile.
Dixit and Padmadas [32]	– The help system can guide the user by offering link suggestion to followed next corresponding to user goal.	– Limited to learnability and navigability – Conducted in virtual laboratory

Table 2. Automated UX evaluation strengths and weaknesses *(Continued)*

Author(s)	Strengths	Weaknesses
Scheller and Kühn [33]	– The proposed framework is extendable to add new UX attributes and metrics for further experiments	– Cannot measure the naming and placement of API elements that fits the expectations of the user. – Complexity of the documentation is not explicitly taken into account yet.
Feijó et al. [10]	– Proposed system have a unique feature to collect and log emotions in automated UX evaluation	– Limited to cognitive workload and expressive reaction
Harms, [34]	– The proposed approach is fully automated. – The approach can easily be practiced by individual who do not have an extensive expert of automated UX evaluation in general.	– Unused GUI Elements cannot be automated. – The approach recorded all user events and do not record only events that represent user action instances which can lead to invalid results of the task generation and the UX detection
Ferre et al. [35]	– The extended approach can be conducted in UX evaluation lab, automating the quantitative metrics, and for logging real application in use.	– Did not determine UX attributes and related metrics in capturing phase
Moran et al. [36]	– The proposed approach can detect mobile interface design violations – The design violations developed based on robust and in use industrial dataset.	– Identify common errors in the implementations of mobile GUIs and neglects the human factor
Mathur et al. [37]	– The framework depends on the functional UX assessment by considering mobile UX guidelines and code analysis. – Provide an code enhancement recommendation in case of UX failure	– Evaluate the functional aspects of UX and less on human aspects
Souza and Canedo [6]	– The adopted questionnaire (System UX Post-Study questionnaire) that have been automated has a high reliability coefficient	– No validation was conducted to check the validity of automated tool – Limited to subjective (questionnaire) automation
Ntoa et al. [7]	– The proposed tool does not require any instrumentation, in contrast to previous efforts in the field of UX evaluation automation.	– Does not support the subjective evaluation
Sodhar et al. [42]	– The study can be generalized since the tool is freely available	– General Attributes – Does not mention any UX metric
Saleh et al. [43]	– Collect UX metrics without user observation (automated gathering)	– Need instrumentation to measure ux logging metric – Increase page load time

In fact, existing UX evaluations in literature so far lack empirical evidence, particularly with respect to selection of empirical methods for UX evaluation. Main observation from this review is that there is an essential need for systematic mapping between UX attributes and UX evaluation methods, which can minimize intricacies and enhance results of evaluation process. In general, a number of different UX attributes and factors must be considered when developing tools and frameworks for UX evaluation.

In this context, automated UX evaluation can be both user-reported oriented or automated usage tracking based on UX attributes and metrics determined by UX evaluator. Several UX attributes and metrics can be further added into automated evaluation due to the flexibility of objective and subjective measurement.

6 Conclusion and future work

UX evaluation is consequently to produce and manage effectively and efficiently usable mobile application, which not only represent real needs of users but also satisfy their expectations. Primary subject of this study is automated UX evaluation of mobile application, which has been identified as UX evaluation tools, where strengths and weaknesses of introduced work are identified and analyzed. Moreover, a variety of automated UX tools, UX evaluation studies together with their strengths and weaknesses have been discussed.

In this study, we provided an overview of automated UX evaluation methods for mobile application, which are often performed to measure UX metrics that some methods neglect. Automated tools can identify whether the user is satisfied or not (it reflects actual user behavior and what users are doing), but cannot identify the reason behind this observation (why users are doing this and their intentions; the reason behind the action). Automated tools are good for gathering specific insights about peoples' actual behavior' unlike surveys since if you base your design decisions on what people tell you they do, and not what they actually do, your product will not be designed to support actual user behavior.

In light of the literature review summary of the state-of-the-art and considering the strengths and weaknesses reported in the literature, there are future opportunities for

1. Developing new automated UX tools for mobile applications that address the issues that have been highlighted from the UX evaluation. This could improve the UX of mobile application and increase the user experience and satisfaction in using them. These tools should be designed to cover different scenarios such as redesigning an existing product or designing new products and services or to use in conjunction with other evaluation methods.
2. Exploring UX evaluation attributes, which fit the mobile features and factors, such as use, task and context in matter of automation the subjective measurements specifically questionnaire automation. This will contribute to extend the automated UX evaluation from the objective measurement to fill the gap in the current automated UX evaluation.
3. Several UX attributes and metrics can be further added into automated evaluation due to the flexibility of objective and subjective measurement.

However, it was appealingly observed that simplicity and cognitive load attributes were rare and only used in UX evaluations that are involved in mobile application contexts. Thus, it is considering as motivation to involve those two attributes in future work for its importance as discussed earlier in this study.

As matter of fact, the automated UX evaluation might be conducted within development process and should not be considered as unique evaluation method and it has to be complemented with traditional UX evaluation, such as CW, observation, think aloud and heuristic evaluation.

7   References

[1] B. Zaid, R. Jamaludin, and B. Wafaa. A comparative study of usability methods for mobile applications. International Journal of Scientific & Engineering Research, vol. 3, no. 8, pp. 1–4, 2012.

[2] K. L. Norman and E. Panizzi, "Levels of automation and user participation in usability testing." Interacting with Computers 18, in University of Maryland, Laboratory for, 2006: Citeseer. https://doi.org/10.1016/j.intcom.2005.06.002

[3] A. Hussain, H. A. Razak, and E. O. Mkpojiogu, "The perceived usability of automated testing tools for mobile applications," Journal of Engineering, Science and Technology (JESTEC), 2017.

[4] C. Coursaris and D. Kim, "A qualitative review of empirical mobile usability studies," AMCIS 2006 Proceedings, p. 352, 2006.

[5] M. de Sá and L. Carriço, "Defining scenarios for mobile design and evaluation," in CHI'08 extended abstracts on human factors in computing systems, ACM, pp. 2847–2852, 2008. https://doi.org/10.1145/1358628.1358772

[6] P. G. d. S. e Souza and E. D. Canedo, "Improving usability evaluation by automating a standardized usability questionnaire," in International Conference of Design, User Experience, and Usability, Springer, 2018, pp. 379–395, 2018. https://doi.org/10.1007/978-3-319-91797-9_27

[7] S. Ntoa, G. Margetis, M. Antona, and C. Stephanidis, "UXAmI observer: An automated user experience evaluation tool for ambient intelligence environments," in Proceedings of SAI Intelligent Systems Conference, Springer, 2018, pp. 1350–1370, 2018. https://doi.org/10.1007/978-3-030-01054-6_94

[8] A. Oztoprak and C. Erbug, "Field versus laboratory usability testing: A first comparison," Tech. rep., Department of Industrial Design-Middle East Technical University, 2008.

[9] F. Paz and J. A. Pow-Sang, "A systematic mapping review of usability evaluation methods for software development process," International Journal of Software Engineering and Its Applications, vol. 10, no. 1, pp. 165–178, 2016. https://doi.org/10.14257/ijseia.2016.10.1.16

[10] J. Feijó Filho, W. Prata, and T. Valle, "Emotions logging in automated usability tests for mobile devices," in International Conference of Design, User Experience, and Usability, Springer, pp. 428–435, 2015. https://doi.org/10.1007/978-3-319-20886-2_40

[11] A. Bruun and J. Stage, "An empirical study of the effects of three think-aloud protocols on identification of usability problems," in Human-Computer Interaction, Springer, pp. 159–176, 2015. https://doi.org/10.1007/978-3-319-22668-2_14

[12] J. Nielsen, Usability engineering. Elsevier, 1994. https://doi.org/10.1016/B978-0-08-052029-2.50007-3

[13] A. W. Kushniruk, H. Monkman, D. Tuden, P. Bellwood, and E. M. Borycki, "Integrating heuristic evaluation with cognitive walkthrough: Development of a hybrid usability inspection method," Studies in Health Technology and Informatics, vol. 208, pp. 221–225, 2015.

[14] Y. Liu, A.-L. Osvalder, and S. Dahlman, "Exploring user background settings in cognitive walkthrough evaluation of medical prototype interfaces: A case study," International Journal of Industrial Ergonomics, vol. 35, no. 4, pp. 379–390, 2005. https://doi.org/10.1016/j.ergon.2004.10.004

[15] J.-M. López-Gil, M. Urretavizcaya, B. Losada, and I. Fernández-Castro, "Integrating field studies in agile development to evaluate usability on context dependant mobile applications," in Proceedings of the XV International Conference on Human Computer Interaction, ACM, p. 24, 2014. https://doi.org/10.1145/2662253.2662277

[16] A. Kaikkonen, A. Kekäläinen, M. Cankar, T. Kallio, and A. Kankainen, "Usability testing of mobile applications: A comparison between laboratory and field testing," Journal of Usability Studies, vol. 1, no. 1, pp. 4–16, 2005.

[17] F. Balagtas-Fernandez and H. Hussmann, "A methodology and framework to simplify usability analysis of mobile applications," in Proceedings of the 2009 IEEE/ACM International Conference on Automated Software Engineering, IEEE Computer Society, pp. 520–524, 2009. https://doi.org/10.1109/ASE.2009.12

[18] J. Xu, X. Ding, K. Huang, and G. Chen, "A pilot study of an inspection framework for automated usability guideline reviews of mobile health applications," in Proceedings of the Wireless Health 2014 on National Institutes of Health, ACM, pp. 1–8, 2014. https://doi.org/10.1145/2668883.2669585

[19] W. Kluth, K.-H. Krempels, and C. Samsel, "Automated usability testing for mobile applications," in WEBIST (2), pp. 149–156, 2014.

[20] A. Dingli and J. Mifsud, "Useful: A framework to mainstream web site usability through automated evaluation," International Journal of Human Computer Interaction (IJHCI), vol. 2, no. 1, p. 10, 2011.

[21] R. Harrison, D. Flood, and D. Duce, "Usability of mobile applications: Literature review and rationale for a new usability model," Journal of Interaction Science, vol. 1, no. 1, p. 1, 2013. https://doi.org/10.1186/2194-0827-1-1

[22] ISO 9241-11. Ergonomic requirements for office work with Visual Display Terminals (VDTs), 1998.

[23] X. Ma, B. Yan, G. Chen, C. Zhang, K. Huang, and J. Drury, "A toolkit for usability testing of mobile applications," in International Conference on Mobile Computing, Applications, and Services, Springer, pp. 226–245, 2011. https://doi.org/10.1007/978-3-642-32320-1_15

[24] A. H. Kronbauer, C. A. Santos, and V. Vieira, "Smartphone applications usability evaluation: A hybrid model and its implementation," in International Conference on Human-Centred Software Engineering, Springer, pp. 146–163, 2012. https://doi.org/10.1007/978-3-642-34347-6_9

[25] D. Bader and D. Pagano, "Towards automated detection of mobile usability issues," in Software Engineering (Workshops), pp. 341–354, 2013.

[26] F. T. Au, S. Baker, I. Warren, and G. Dobbie, "Automated usability testing framework," in Proceedings of the Ninth Conference on Australasian User Interface-Volume 76, Australian Computer Society, Inc., pp. 55–64, 2008.

[27] A. Følstad, E. Law, and K. Hornbæk, "Analysis in practical usability evaluation: a survey study," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, ACM, pp. 2127–2136, 2012. https://doi.org/10.1145/2207676.2208365

[28] K. Kiili, "Evaluating WAP usability: "What usability?"," in Wireless and Mobile Technologies in Education, 2002. Proceedings. IEEE International Workshop on, pp. 169–170, 2002.

[29] H. B.-L. Duh, G. C. Tan, and V. H.-h. Chen, "Usability evaluation for mobile device: a comparison of laboratory and field tests," in Proceedings of the 8th Conference on Human-Computer Interaction with Mobile Devices and Services, ACM, pp. 181–186, 2006.

[30] M. Zduniak, "Automated gui testing of mobile java applications," Master's thesis, Poznan University of Technology Faculty of Computer Science and Management, 2007.

[31] F. Lettner and C. Holzmann, "Automated and unsupervised user interaction logging as basis for usability evaluation of mobile applications," in Proceedings of the 10th International Conference on Advances in Mobile Computing & Multimedia, ACM, pp. 118–127, 2012. https://doi.org/10.1145/2428955.2428983

[32] S. Dixit and V. Padmadas, "Automated usability evaluation of web applications," in Proceedings of the International Congress on Information and Communication Technology, Springer, pp. 139–149, 2016. https://doi.org/10.1007/978-981-10-0755-2_16

[33] T. Scheller and E. Kühn, "Automated measurement of API usability: The API concepts framework," Information and Software Technology, vol. 61, pp. 145–162, 2015. https://doi.org/10.1016/j.infsof.2015.01.009

[34] P. Harms, "Automated field usability evaluation using generated task trees," 2016.

[35] X. Ferre, E. Villalba, H. Julio, and H. Zhu, "Extending mobile app analytics for usability test logging," in IFIP Conference on Human-Computer Interaction, Springer, pp. 114–131, 2017. https://doi.org/10.1007/978-3-319-67687-6_9

[36] K. Moran, B. Li, C. Bernal-Cárdenas, D. Jelf, and D. Poshyvanyk, "Automated reporting of GUI design violations for mobile apps," arXiv preprint arXiv:1802.04732, 2018. https://doi.org/10.1145/3180155.3180246

[37] N. Mathur, S. A. Karre, and Y. R. Reddy, "Usability evaluation framework for mobile apps using code analysis," in Proceedings of the 22nd International Conference on Evaluation and Assessment in Software Engineering 2018, ACM, pp. 187–192, 2018. https://doi.org/10.1145/3210459.3210480

[38] R. Patton, Software testing. Pearson Education India, 2006.

[39] ISO, "9241-11: 1998," Ergonomic requirements for office work with visual display terminals (VDTs)–Part II guidance on usability, 1998.

[40] H. Y. Abuaddous, M. Z. Jali, and N. Basir, (2017). Quantitative metric for ranking web accessibility barriers based on their severity. Journal of Information and Communication Technology, 16(1), 81–102. https://doi.org/10.32890/jict2017.16.1.8219

[41] H. Y. Abuaddous, M. Z. Jali, and N. Basir, "Web accessibility challenges," International Journal of Advanced Computer Science and Applications, vol. 7, no. 10, pp. 172–181, 2016. https://doi.org/10.14569/IJACSA.2016.071023

[42] I. N. Sodhar, A. A. Mirani, and A. N. Sodhar, "Automated usability evaluation of government and private sector educational websites of pakistan," Information Sciences Letters, vol. 8, no. 2, pp. 51–55, 2019. https://doi.org/10.18576/isl/080202

[43] A. M. Saleh, A. B. Al-Badareen, and O. Enaizan, "Automated user experience tool development for mobile application," International Journal of Advanced Science and Technology, vol. 29, no. 6, pp. 4929–4940, 2020.

[44] B. Richardson, M. Campbell-Yeo, and M. Smit, "Mobile application user experience checklist: A tool to assess attention to core UX principles," International Journal of Human–Computer Interaction, pp. 1–8, 2021. https://doi.org/10.1080/10447318.2021.1876361

[45] Dollars and Sense: The business case for investing in UI design. Accessed: 5th July, 2021.

[46] D. Albers, J. Radlmayr, A. Loew, S. Hergeth, F. Naujoks, A. Keinath, and K. Bengler, "Usability evaluation—advances in experimental design in the context of automated driving human–machine interfaces," Information, vol. 11, no. 5, p. 240, 2020. https://doi.org/10.3390/info11050240

[47] S.-Y. Chien, Z. Semnani-Azad, M. Lewis, and K. Sycara, "Towards the development of an inter-cultural scale to measure trust in automation," In Cross-Cultural Design; Rau, P.L.R., Ed.; Springer International Publishing: Cham, Switzerland, 2014; ISBN 978-3-319-07307-1.

[48] A. M. Saleh, H. Y. Abuaddous, O. Enaizan, and F. Ghabban, "User experience assessment of a COVID-19 tracking mobile application (AMAN) in Jordan," Indonesian Journal of Electrical Engineering and Computer Science, vol. 23, no. 2, pp. 1120–1127, 2021. https://doi.org/10.11591/ijeecs.v23.i2.pp1120-1127

8   Authors

Hayfa. Y. Abuaddous, Asst. Prof, Software Engineering Department, Faculty of Computer Sciences and Informatics, Amman Arab University, Jordan, E-mail: haddose@aau.edu.jo

Ashraf Mousa Saleh, Asst. Prof, Software Engineering Department, Faculty of Computer Sciences and Informatics, Amman Arab University, Jordan, E-mail: asaleh@aau.edu.jo

Odai Enaizan, Asst. Prof,Haql University College, Department of Management Information System, University of Tabuk, Saudi Arabia, E-mail: aonaizan@ut.edu.sa

Fahad Ghabban, Asst Prof, Information system, College of Computer Science and Engineering, Taibah University, Saudi Arabia. Email: fghaban@taibahu.edu.sa

Anas Bassam Al-Badareen, Assoc. Prof, Department of Software Engineering, Aqaba University of Technology, Jordan Email: abdareen@aut.edu.jo