A Methodology for Verification of a New Model for Evaluating the Usability of an M-banking Application

Hamisu Ibrahim Abubakar[1], Nor laily Hashim[2], and Azham Hussain[3]

***Abstract*—**Many models for evaluating the usability of mobile application exists, but they are static, crumbled and have not undergone a verification process by usability experts to examine their capability of collecting data for the intended applications. Therefore, the proposed model is established and verified in response to the need identified in the literature because it is important for m-banking applications to provide customers with the anticipated and likely sense of interaction, to be easy to use and encourage the customers to accept the technology. This paper describes the verification procedure for a newly proposed model for evaluating the usability of m-banking applications. The purpose of this verification procedure is to identify the main practices for model verification methods for evaluating the usability of an m-banking application. The verification was conducted through the use of usability experts in mobile application development and academia to examine the model and its components. The verification form and questionnaire that measured the model in terms of consistency, understandable, easy to use, tailorable, verifiable and overall impressions have been completed by the usability experts, and the proposed model has been improved based on the feedback received from the experts. The finding from the experts' questionnaire shows that the proposed model is complete, original and acceptable for the intended application. Therefore, this paper will provide additional knowledge in both theory and practice on model verification methods, especially for usability evaluations of commercial applications.**

***Keywords*—**Verification process, Expert review, Usability evaluation, M-banking application.

I. Introduction

Mobile banking (M-banking) can be described as the banking transactions and services that a user can perform via a mobile device at any time and from any place at the user's convenience. Most m-banking applications have the same functionalities. They provide a variety of financial transactions, such as bill payments, fund transfer, recharging of a card, investment and insurance. Evaluating the usability of an m-banking application is important because the designers and developers can thereby identify the strengths and weaknesses of the applications [16], [8]. Similarly, it captures the efficiency, effectiveness and accuracy of the application. However, in order to achieve usability evaluation for an m-banking application, there is a need for a comprehensive usability evaluation approach containing appropriate usability measurements [9], [1].

[1,2,3], School of Computing, Universiti Utara Malaysia, 06010, Sintok, Kedah, Darul Aman. E-mail: H.I. Abubakar, haibalgnm@gmail.com; Nor laily, H., laily@uum.edu.my; Azham, H., azham.h@uum.edu.my.

Many usability evaluation models have been proposed, such as ISO 9241-11, ISO 9126-1, and mGQM [13], [2], [3]. However, most of these models are based on design and are not specific to a particular mobile application [8], [2]. Additionally, such models have not been sufficiently used in evaluating the usability of m-banking applications. Consequently, these models do not provide overall descriptions of how to select metrics corresponding to the usability factors or criteria [24].

Therefore, a new usability evaluation model for m-banking applications is built through the establishment of usability factors, criteria and metrics. However, the identified components of the developed model are within the four usability contextual factors: user, environment, technology and task [7], [2]. Consequently, these usability measurements are seen as sufficient for use in evaluating the usability of an m-banking application because they are comprehensive and carefully support Human Computer Interaction (HCI) principles [2], [16], [21].

The main part of this paper is devoted to explaining the model verification method. The model itself contributes towards a better understanding of the modelling usability approach. Therefore, the model has now reached the final stage of development and the need to verify the model components becomes imperative.

II. Problem Formulation

Many studies on mobile application usability approaches concentrate mainly on one aspect of the evaluation without addressing a proper model verification approach [4], [26], [27]. Therefore, the model may not provide a satisfactory range of accuracy for a validation process [26]. Consequently, little literature exists that relates to how the model and its components have been verified by experts [4], [11].

III. Proposed Usability Evaluation Model for M-banking Application

The proposed usability evaluation model consists of three key components that include usability factors, criteria and metrics. The usability factors consist of efficiency, effectiveness, user satisfaction, learnability and trustfulness [1], [2], [5], [6], [7], [20], as shown in Table 1. These five usability factors are derived based on prioritization in the reviewed literature, importance and relevance to m-banking functionalities. The derived usability factors support task/activities and technology (device) contexts.

The usability factors are broken down into measurable

criteria (sub-factors). A criterion is directly measurable via a particular metric that is linked to a usability factor [20], [10], [12]. In other words, the measurement of a usability factor depends on the corresponding criteria, while metrics are used to measure criteria that are assigned to a particular usability factor. Metrics are defined in terms of formulas or countable metrics that are extracted from raw data such as video observation or experiment, depending on the application type [10], [15]. The usability metrics are classified into two main categories: testing and predictive metrics [22], [8], [25]. The testing metrics are used to collect data to measure the actual use or function of the working application and identify problems associated with the application [7], [20]. In collecting this type of data, a fully functional application or prototype is required. However, preference metrics deal with the subjective evaluation level of end users' satisfaction and performance metrics measure the real performance of the users while accomplishing a task [25]; all are classified under testing metrics. Therefore, the developed model is designed based on testing metrics that consist of preference and performance metrics.

Fourteen criteria have been generated through the use of the Systematic Literature Review [22], [7]. Each criterion is positioned to its corresponding usability factors based on support in the literature. These groupings of criteria to usability factors have been used and agreed to by many usability evaluations papers, both in mobile applications and the software domain, such as [20], [14], [24], [9].

Furthermore, to generate considerable usability metrics for the generated criteria, the defined metrics in GQM [3], the QUIM model [20] and other usability studies such as mGQM [13], [14] were critically analysed and employed. Therefore, 43 metrics are derived both for subjective and objective measurements. However, 17 measure objective data and 26 measure subjective measurements. Table 1 below shows the grouping of criteria to their corresponding factors, as does Fig. 1 in Appendix 1 for the first version of the proposed model of the m-banking usability evaluation.

Table 1: Grouping of criteria to their corresponding usability factors

Criteria	**Usability factor**				
	Efficiency	Effectiveness	Trustfulness	Learnability	User satisfaction
Compatibility	√				
Loading time	√				
Operability	√				
Accuracy		√			
Presentation		√			
Navigation		√			
Privacy			√		
Reliability			√		
Simplicity				√	
Familiarity				√	
Consistency				√	
Content					√
Structured task					√
User guide					√

IV. VERIFICATION METHODOLOGY

This activity was performed as a first evaluation of the developed model. Moreover, the verification process determined whether all of the metrics should remain in the proposed model or whether some needed to be removed. Verification is a technique for ensuring that the model and the components, as well as other entities within the model, are sufficient, accurate and complete for its purpose or to determine whether the model is being built using an orderly approach [4], [18]. However, verification is used to confirm that all of the components of the model possess a satisfactory range of accuracy and completeness and are consistent with the intended application.

A. Experts Review

The experts' review examines the capability of collecting data for each metric. [5], as cited in (Lauesen and Vinter, 2001), states that the reliability of using expert decision is very high when put into practice. Similarly, integrating experts in both theory and practice can lead to simple and accurate results [12], [4]. Usability experts provide speedy and valuable comments that will improve the quality of the model's design and development [17]. The expert may be independent from the development team and willing to give honest opinions and comments [23].

B. Purpose of the Verification Method

The goal of the verification method is to obtain feedback from usability experts regarding the originality, reliability and completeness of the proposed model [4]. Furthermore, the expert review and verification is used to provide empirical evidence from the field of HCI for academia and m-banking application developers in the industry. It will also enhance originality, richness and quality flow of the measurements [24].

C. Instrument Development

To provide a good verification approach, experts' review documents, experts' verification forms and a questionnaire were designed as instruments for the expert reviewers. The experts' review document consists of details of the process of model development, five segments of the proposed model, a list of usability factors (and their measurements), data calculation methods for objective metrics, a list of tasks for the experiment and the designed model. Moreover, the verification instruments contain three sections: section A is the expert's profile, section B contains a set of metrics with corresponding criteria and usability factors, and section C is the questionnaire. The questionnaire came from [4] with little modification, and it contains five measurable factors with two (2) scale options: "Agree" or "Disagree". The measuring factors used to build the questionnaire include the following: 1) Consistency, 2) Understandable, 3) Ease of use, 4) Tailorable and 5) Verifiable. However, to measure the expert's opinion regarding the model, Overall impression was also included in the questionnaire. The six dimensions are used by the expert to rate the model in terms of its originality and acceptability for use by usability practitioners and for research purposes [4]. The verification form contains details regarding the model, which includes components of the model and the relationships that exist between each

entity and the six dimensions in the experts' questionnaire which the model is judged. Additionally, items included in the experts' verification documents are objectives of the model and the instrument for model testing. Table 2 below presents the data calculation methods for the objective metrics.

Table 2. Data calculation method of objective metrics

Metrics	**Calculation method**
Time taken to load/initialize	Finish load/initialize time-start loading time
Time taken to connect to the network	Connection time – start connect to network time.
Number of interactions while keying-in user ID and Password	Count how much time participant has to interact with the application during keying-in user ID & password
Time taken to respond	Finish processing time - start processing time
Time taken to display transaction page	Display transaction page time - time for main menu to appear.
Time taken to start transaction	Time for main menu to appear - time participant completed keying-in user ID and Password.
Time taken to select a task	Time of selection of a task - start selection time.
Time taken to finish a task	Time to finish the task – time started the task.
Number of error (s) during a task	Count the errors made by participant for each task.
Time taken to navigate	Time to start a task -navigation time
Number of steps during task selection	Count how many steps for participant to select a task
Number of Interactions while performing a task	Count how much time participants have to interact with the application during a given task.
Time taken to display output	Output display time - task finishing time.

D. Data Collection

The experts were contacted, and discussion has occurred concerning the improvement of the proposed model. Six experts were selected from academic institutions and the m-banking application developers' industry. [19] mentioned that three to five experts are sufficient to review and verify a newly developed model.

To achieve an effective model verification process [18], the designed experts' instruments were distributed to the selected usability experts in both academia and industry. Four of the selected experts were contacted via email followed by telephone calls, while two experts were contacted directly in their respective offices to obtain additional comments and suggestions.

E. Data Analysis

Data collected from the experts were sorted and stored in the SPSS statistical package for analysis. SPSS version 22 was used to find the means of individual measurements in the experts' instrument. The results were compared and analysed based on individual expert scores to examine the substantiality of the result. The result of the data collected from the experts is presented in Table 7. All of the comments and suggestions received from the experts were carefully examined and considered for improvement of the proposed model. The unrelated and repeated metrics based on the experts' comments/suggestions were critically and objectively analysed.

V.RESULT/DISCUSSION

The experts were generally helpful in this verification process, with general agreement that it will pinpoint flaws in the proposed model and assist the researcher with improvement activities. The verification provides a positive experience and has provided some significant contributions to the developed model. Based on the suggestions and comments made by the experts, it appears that some metrics are repeated, while a few are not relevant to the intended application. The unrelated and repeated metrics have been removed, to avoid obstacles for the proposed model. Experts commented that the definition of the usability factors, criteria and many of the metrics are relevant and capable of collecting data. Additionally, the experts agreed that the proposed model is specific to the m-banking application and can be expected to provide satisfactory results when tested. However, the experts' comments and suggestions made a significant contribution to the proposed model. The consultations with the experts provided a positive experience and have provided some important contributions for the development of the proposed model. Please refer to Table 3 for overall comments given by each expert.

Table 3. Comments/suggestions received from the experts

Expert	**Comments/Suggestions**
A	The metrics are good for the evaluation of usability in m-banking applications, but the experts have highlighted the importance of privacy and user guides. Metrics such as "single sign on", "session timeout", "alert/warning message", and authentication should be included.
B	The metrics are relevant to the defined usability criteria, but the metric "satisfaction with help" need not be overemphasized. Metric "time taken to respond" should be modified to "Response time", and "menu name" should be changed to "menu item", and "linked list" can be modified to "sub-menus". Performance speed should be considered.
C	The metrics are generally suitable and relevant to the defined criteria in the proposed model.
D	The experts suggested that metrics such as "Number of errors during key-in user detail" should be under "Accuracy", criteria, and "Number of interactions" should be modified to "Number of interactions per unit time". The word "attempt" can be modified to "step". The metric "Time taken to navigate a menu" can be included under "Navigation" criteria.
E	Some mobile devices do not provide a keypad anymore. The metrics are suitable, but more aspects need to be explored in each dimension. More metrics should be added under "Navigation" criteria.
F	All the features are relevant and important. Questions on Overall Satisfaction and "Satisfaction of the task" as metrics should be included. Tasks should be selected for the evaluation, and the objective of the tasks needs to be explained

Based on the experts' comments, metrics that are not related to the defined criteria have been removed, while additional metrics suggested by the experts were added to

their corresponding criteria for improvement of the model, as shown in Tables 5 and 6.

Experts were asked to judge the proposed model using six dimensions, namely consistency, understandable, easy to use, tailorable, verifiable and overall impression. These dimensions measured the originality, completeness and acceptability of the developed model. Two options were given in the questionnaire instrument, "Agree" and "Disagree". The result shows that all experts tended to choose "Agree". Please refer to Table 7 for mean scores of the individual experts.

Table 4. Experts' Profiles

Expert	Profile
A	Specialization: Mobile application developer Position: Senior Software Engineer Year of experience: 4 years.
B	Specialization: Software Engineering Position: Senior Lecturer Year of experience: 14 years.
C	Specialization: Mobile application developer Position: Director General Year of experience: 8 years.
D	Specialization: Interaction Design Position: Senior Lecturer Year of experience: 14 years.
E	Specialization: HCI/Interaction Design Position: Associate Professor Year of experience: 14 years.
F	Specialization: HCI & Software Engineering Position: Senior Lecturer Year of experience: 14 years.

Table 4 above described the usability experts' profiles, where two (2) experts are from industry and four (4) experts are from academia.

Table 5. List of dropped metrics

Dropped metrics	Suggested by expert
Number of clicks to sign-in	A, C, D & F
Number of attempts to sign-in/sign-out	A, B, C & D
Satisfaction with help	A, B, C, D, & E
Satisfaction with menu names	A, B, C & E
Time taken to key-in user details	A, B, C, E & F
Satisfaction with device keypad	A, C, E & F
Number of error during key-in user data	A, C, D & F
Number of interactions	A, C, D, E & F
Mobile device support	A, C, D & F
Time taken to learn	A, C, D & F
Satisfaction with link-list	A, D, E & F
Time taken to key-in user data	A, B, C &E

Table 5 presents the list of metrics that were dropped from the proposed model as suggested by the experts. Some of the metrics are not relevant for the intended application, while others may not adequately capture the complexity of the m-banking application; as such, those metrics will remain redundant in the proposed model.

Table 6. List of added metrics

Added Metrics	Suggested by expert
Satisfaction with session timeout if idle	A
Satisfaction with alert message if error occurred	A
Satisfaction with authentication technique	A
Performance speed	B
Time taken to navigate	D & E
Number of steps during task selection	D
Satisfaction with the task performed	F
Satisfaction with menu items provided	B
Easy to use and learn	A, B & F

Table 6 describes the metrics that are considered to be important for the usability evaluation of m-banking application as recommended by usability experts. Therefore, these metrics have been added to improve the proposed model.

Table 7: Mean scores for individual experts

Expert	Dimension					
	Consistency	Understandable	Easy to use	Tailorable	Verifiable	Overall impression
A	0.50	1	1	1	1	1
B	.75	1	1	1	1	1
C	1	0.50	1	1	1	1
D	.75	1	1	1	1	1
E	.75	1	1	1	1	1
F	1	1	1	1	1	1

Table 7 above shows that four dimensions score relatively high with 0.75 and 1 by individual experts, whereas two dimensions, namely consistency and understandable, scored a 0.50 by expert A and expert C, respectively. As illustrated in Table 8, all of the individual dimensions score relatively high (overall scores). The dimensions "Easy to use", "Tailorable", "Verifiable" and "Overall impression" obtained the highest score (1). Consistency scored 0.79 and Understandable scored 0.92. Therefore, these results reveal that the developed model for evaluating the usability of m-banking application is original, complete and acceptable.

Table 8. Overall scores for individual dimension

DIMENSION	MEANS
Consistency	.79
Understandable	.92
Easy to use	1
Tailorable	1
Verifiable	1
Overall impression	1

Additionally, the findings indicate that four dimensions acquired the highest score, followed by the dimension "Understandable". The "Consistency" dimension got the lowest score (79%). This could be due to the irrelevancy or inconsistency of some metrics identified by the experts. However, those metrics were removed and a few were added based on the comments and suggestions received from the experts.

All of the comments and suggestions from the experts were carefully examined and given much consideration for improvement of the proposed model. Therefore, an amendment has been made concerning the metrics and their corresponding criteria from the first version of the model based on the experts' feedback. For instance, previously, the model contains a total of 49 metrics, of which 30 are subjective data whereas 19 are objective data. Therefore, the revised version of the developed model consists of 42 metrics. Twenty-nine metrics focus on subjective data, and 13 metrics centre on objective data (see Fig. 2 in Appendix 2 for the revised version of the proposed model). The asterisk (*) in the model represents amended metrics.

VI. CONCLUSION

The need for a reliable and appropriate model verification approach will, of course, continue to grow to have confidence in the outcome of the model after implementation/testing in the mobile application industry platforms [11]. The main objective of this model is to guide usability practitioners and m-banking application developers toward relevant processes in evaluating the usability of m-banking applications.

Moreover, this paper has shown how a group of experts was used to verify the developed usability evaluation model for an m-banking application. Some metrics were removed, while a few have been added to the proposed model, based on the feedback received from the experts. Furthermore, all of the individual dimensions used by the experts to judge the model reached a satisfactory level, that is, in terms of consistency, understandable, easy to use, tailorable, verifiable and overall impression. This indicates that the proposed usability evaluation model is original, complete and acceptable.

Experts' comments and suggestions have provided significant contributions to the developed model. The experts' responses in the verification forms and questionnaire noted some potential strengths and weaknesses of the proposed model. However, their comments and suggestions are supportive of the concept that building the model needs credible and experienced expertise from both academia and industry. Therefore, this paper illustrates good verification practice, and it will serve as a guide to the research community, especially because literature on the model verification process for mobile application evaluation is very limited [4]. The verification has justified whether the model reflects the needs of the m-banking application from a usability perspective [18].

In this regard, the verification process, details of the experts' report and questionnaire results are presented in this paper. However, this verification represents the final stage of the first cycle of model development. The future target of this study is to test the developed model through usability experiments involving real m-banking application users. This will examine the capability of each metric, both objective and subjective, for collecting data.

Appendix 1

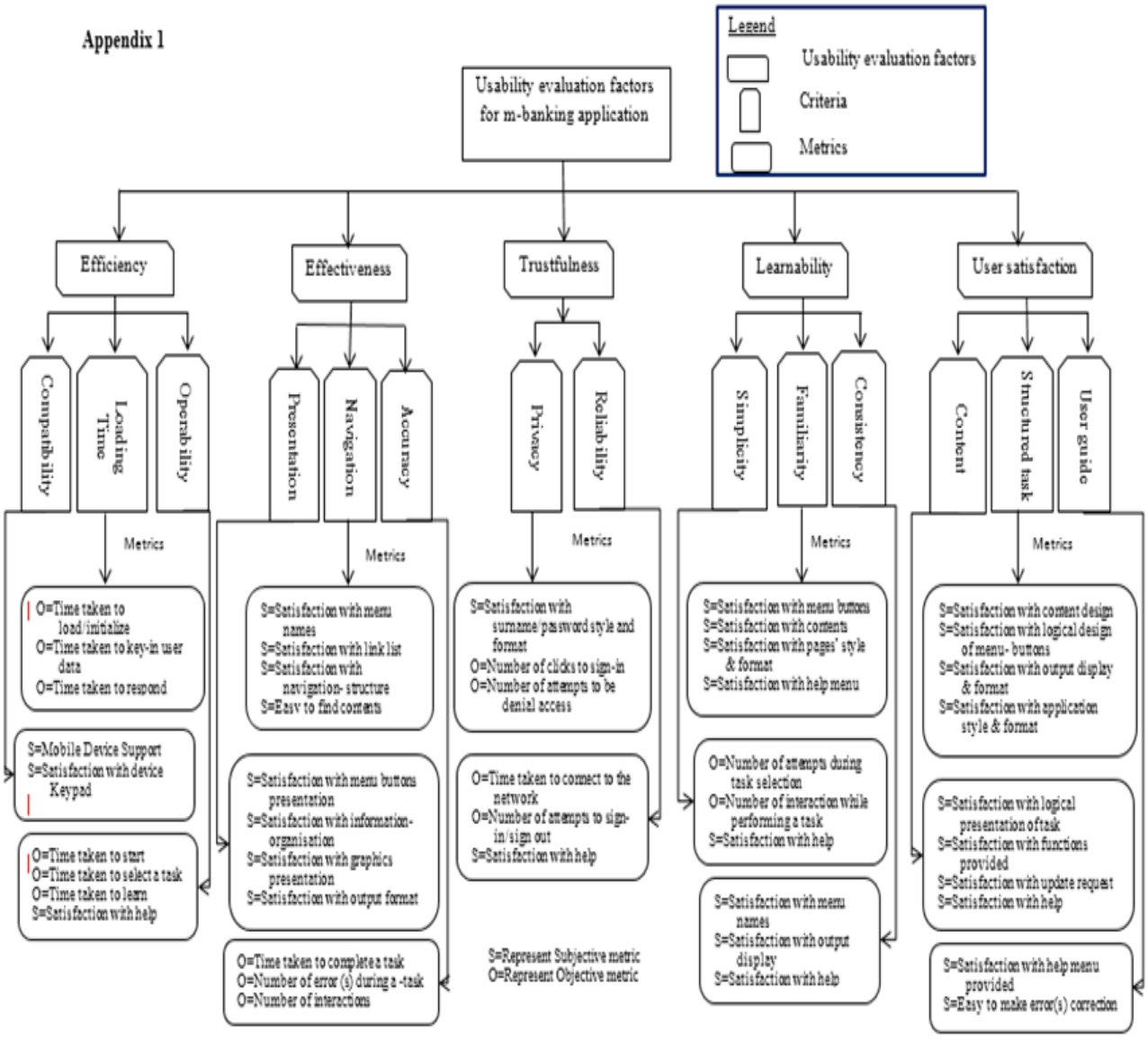

Figure 1. First version of the proposed usability evaluation model for m-banking application

Appendix 2

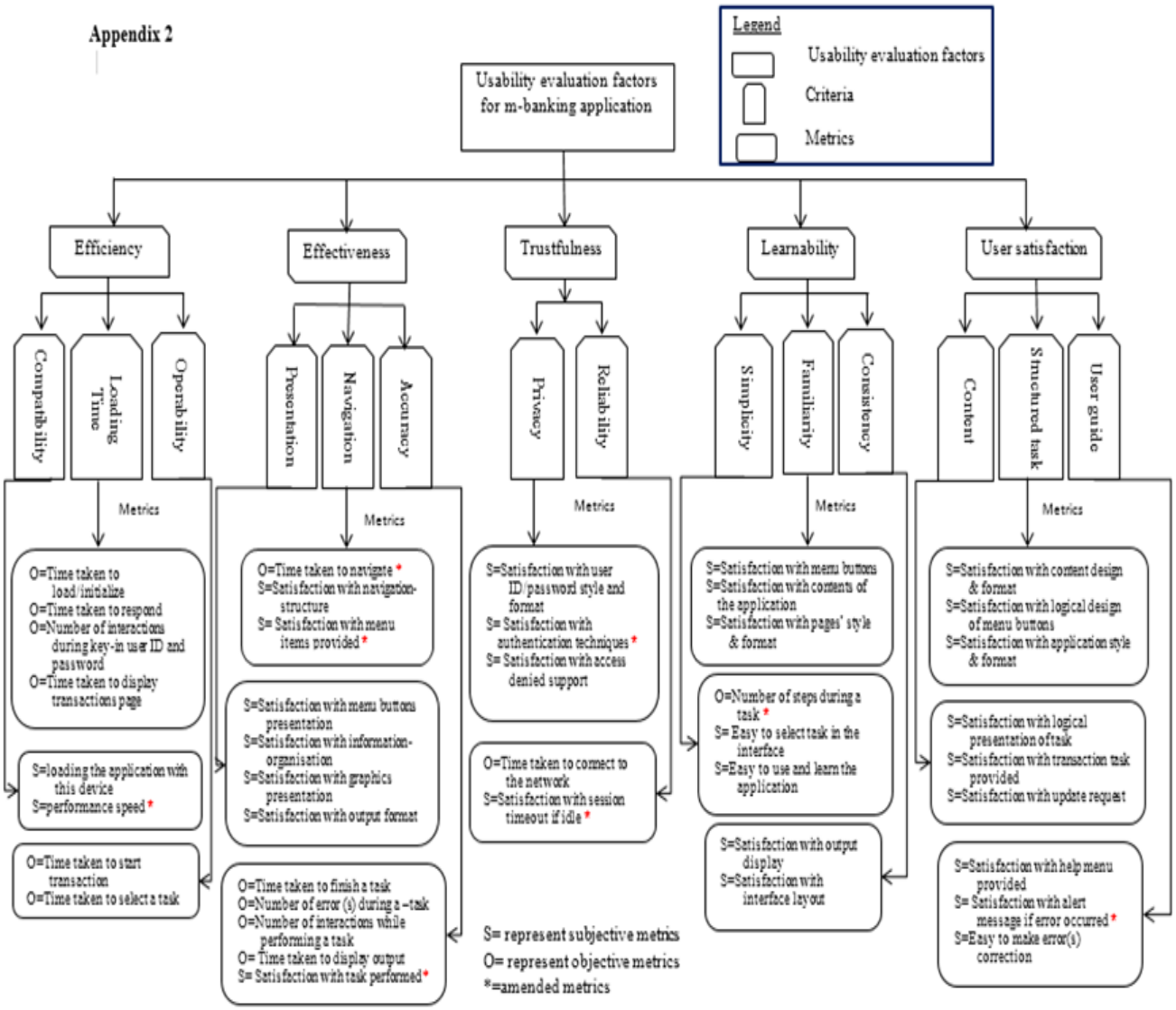

Figure 2. Revised version of the proposed usability evaluation model for m-banking application

REFERENCES

[1] Antunes, P., Herskovic, V., Ochoa, S. F., & Pino, J. A., Structuring dimensions for collaborative systems evaluation. *ACM Computing Surveys (CSUR), 44* (2), 8. 2012.

[2] Baharuddin, R., Singh, D., & Razali, R., Usability Dimensions for Mobile Applications-A Review. *Research Journal of Applied Sciences, 5 (6), p2225-2231*, 2013.

[3] Basili, V. R., Software modeling and measurement: The Goal/Question/Metric paradigm. University of Maryland, College Pack, 1992.

[4] Beecham, S., Hall, T., Britton, C., Cottee, M., & Rainer, A., Using an expert panel to validate a requirements process improvement model. *Journal of Systems and Software, 76*(3), 251-275, 2005.

[5] Bertoa, M. F., Troya, J. M., & Vallecillo, A., Measuring the usability of software components. *Journal of Systems and Software, 79*(3), 427-439, 2006.

[6] Cooharojananone, N., Kongnim, P., Mongkolnut, A., & Hitoshi, O., *Evaluation Study of Usability Factors on Mobile Payment Application on Two Different Service Providers in Thailand.* Paper presented at the 12th International Symposium on Applications and the Internet (SAINT), 2012.

[7] Coursaris, C. K., & Kim, D. J., A meta-analytical review of empirical mobile usability studies. *Journal of Usability Studies, 6*(3), 117-171, 2011.

[8] Dubey, S. K., Gulati, A., & Rana, A., Integrated Model for Software Usability. *International Journal on Computer Science and Engineering, 4*, 429-437, 2012.

[9] Emmanuel, E. A., & Nsung-Nza, M. H., Mobile Commerce Interaction Techniques for African Rural Economy Development: A Case Study for Dwesa, 2012

[10] Frøkjær, E., Hertzum, M., & Hornbæk, K., *Measuring usability: are effectiveness, efficiency, and satisfaction really correlated?* Paper presented at the Proceedings of the SIGCHI conference on Human factors in computing systems, 2000.

[11] Komarkoba, J., Jedlicka, M. & Hub, M., Web-based Geographic Information System and their Usability. Proceeding of the 9th WSEAS International Conference on Applied Computer Science, 2010.

[12] Harrison, R., Flood, D., & Duce, Usability of mobile applications: Literature review and rationale for a new usability model. *Journal of Interaction Science, 1*(1), 1-16, 2013

[13] Hussain, A., *Metric based evaluation of mobile devices: mobile goal question metric (mGQM).* University of Salford, 2012.

[14] Hussain, A., & Kutar, M., Usability Metric Framework for Mobile Phone Application. *PGNet, ISBN*, 978-971, 2009.

[15] Hussain, A., Kutar, M., Mutalib, A. A., & Kamal, F. M.: Modeling Subjective Metrics for Mobile Evaluation. *Journal of Information System Research and Innovation*, ISSN 2289-1358, 2011.

[16] Madan, A., & Dubey, S. K., Usability evaluation Methods: A Literature Review. *International Journal of Engineering Science and Technology, 4*(02), 590-599., 2012.

[17] Nielsen, J., Heuristic evaluation. *Usability inspection methods, 24*, 413, 1994.

[18] Lobo, C.J., Mousalli-Kayat, M. & Rivas. Methodology development and software quality metrics in educational application. Proceeding of the 9th WSEAS International Conference on Telecommunication and Informatics, 2010.

[19] Nielsen, J., & Hackos, J. T., *Usability engineering* (Vol. 125184069): Academic press Boston, 1993.

[20] Seffah, A., Donyaee, M., Kline, R. B., & Padda, H. K., Usability measurement and metrics: A consolidated model. *Software Quality Journal, 14*(2), 159-178, 2006.

[21] Simes-Marques, M., & Nunes, I., Usability of interfaces: INTECH open science, Ergonomics-A System Approach, 2012.

[22] Tan, J., Ronkko, K., & Gencel, C., *A Framework for Software Usability and User Experience Measurement in Mobile Industry*. Paper presented at the Software Measurement and the Eighth International Conference on Software Process and Product Measurement (IWSM-MENSURA), Joint Conference, 2013.

[23] Tory, M., & Moller, T. (2005). Evaluating visualizations: do expert reviews work? *Computer Graphics and Applications, IEEE, 25*(5), 8-11, 2005.

[24] Wiegers, K. E., Seven truths about peer reviews. *Cutter IT Journal, 15*(7), 31-37, 2002.

[25] Mohankumar, M. & Anandkumar, M., Empirical Study on Green and Sustainable Software Engineering. www.wseas.us/e-library/conferences/2015/Dubai.

[26] Hussain, A., Abubakar, H.I. & Hashim. N.B. Evaluating mobile banking application: Usability dimensions and measurements. Paper presented at International Conference on Information Technology and Multimedia (ICIMU), 2014.

[27] Abubakar, H.I., Hashim, N., & Hussain, A. Verification Process of Usability Evaluation Model for M-banking Application. Paper presented at 14th International Conference on Applied Computer and Applied Computational Science (ACACOS'15), 2015