A Quality Assessment Model Based on Usability Metrics for M-Health Applications User Interfaces

Mariam Nosheen¹, Malik Muhammad Saad Missen², Muhammad Abuzar Fahiem³

¹,³Department of Computer Science, Lahore College for Women University, Lahore
²Department of Computer Science & IT, The Islamia University Bahawalpur, Bahawalpur

ABSTRACT

Mobile health (m-Health) applications are gaining popularity for the past few decades. The user interfaces of such applications must be usable and comfortable enough so that they can meet the expectations of their intended users (i.e. doctors, patients etc.). Usability is considered to be one of the key quality factors for measuring the usefulness of such applications which makes usability evaluation a vital task. In this paper, we propose a quality assessment model for evaluating the user interfaces of m-Health applications by reviewing the existing literature and frameworks on the evaluation of user interface designs. The proposed assessment model serves as a base for comprehensive usability evaluation consisting of major usability goals, subfactors, and metrics. The effectiveness and reliability of the proposed model is measured through two android m-health applications. The results gathered from usability testing and statistical techniques validated that the model is applicable for the evaluation of the m-Health applications and act as a helping tool for evaluators and user interface designers.

Keywords: Quality, Usability, m-Health, Usability Metrics, User Interfaces

INTRODUCTION

Mobile devices, especially smartphones, are the most popular mobile technology, people are using in their daily life. According to the worldwide statistics, one out of five people has a smartphone. The Pew statistics and m-Health Market report also indicated that in the USA more than 50% people use this technology to get information about their daily health conditions. Today, many healthcare organizations and different mobile manufacturing companies using this technology as a tool and incorporating health-related features in there devices. These m-Health applications cover variety of diseases like blindness, deafness, Parkin, son’s disease, skin cancer, trauma care, pain management, recording sleep, fitness monitoring and so on for their intended users (patients).

The ease of use while interacting with such applications leads the mobile application designers to integrate usability in their design process so that the usage of such applications will become versatile, unique, user friendly and successful.
It was also found that without the proper assessment of any application, its usability cannot be achieved. A number of usability evaluation models and frameworks are available to measure the usability of any software\textsuperscript{15}, but a limited work is found on mobile applications especially m-Health applications\textsuperscript{16}. The existing model's explorer usability in one or two dimensions and rest are left for future use. These models are not designed to evaluate the usability features specific to the mobile m-Health applications. Majority of these models are not validated for real-time applications and they also did not have the ability to extend to the other domains\textsuperscript{16}. Moreover, very limited guidelines are available which relates the usability criteria and the definitions in a systematic way. To overcome these research gaps an extensive literature review is presented in this paper; which synthesize the usability Model, definitions, goals, attributes, and metrics for m-Health applications. The next section of this paper presents the background of the existing usability models, frameworks. Section 3 describes the research methodology. Section 4 explains the proposed assessment model. Section 5 presents the results and the last section provides the conclusion and future enhancement of the proposed research work.

## BACKGROUND

### Software Quality and Usability

The role of quality in the development of software products was found in the late 60's. Companies incorporate this factor in their production, manufacturing and designing process to increase their product's acceptance and revenue. The rapid technological advancements, socialization, and marketing strategies play an influential role in enhancing the quality of any product. Many standard bodies like ISO/IEC 9126 define generic quality standards for any product\textsuperscript{17,18}. These quality standards divide the quality into many factors like functionality, reliability, usability, efficiency, maintainability, portability\textsuperscript{19}. Usability is different from other factors as it directly deals with the human interaction issues related to the product. In literature, usability is also known as “quality in use” concept\textsuperscript{20}. Where the quality of any product is assessed through usage and interaction. In the software industry, usability deals with the designing issues of interactive systems\textsuperscript{21}. The basic purpose of designing interactive systems is to meet the user expectations and demands. In the software industry, users’ needs and expectations towards the systems are expressed in the form of requirements which are then measured using metrics\textsuperscript{20}. Without the proper measurement, designers cannot find the interaction problems of their system. Various researchers classify the usability of multiple genetic factors. The description of various usability models and factors is discussed in the next section.

#### Usability Models

The definition of usability evolves over a past few decades and its concept has been defined in various ways. In literature, researchers classify the concept of usability into many factors and define various usability models which are generic by nature and unable to cover all the aspects of usability.

A number of usability models are available such as MUSiC\textsuperscript{22} which provides the metrics to evaluate the usability of software applications. The Software Usability Measurement Inventory (SUMI) model provides measures regarding global satisfaction of five specific usability areas\textsuperscript{23}. McCall divided its model into three criteria (i.e. training, operability, and communication)\textsuperscript{24}. The semi-automated interface designer evaluator (AIDE) is used for measuring the static web pages based on pre-defined guidelines\textsuperscript{25}. In 1996, John and Kiers proposed a model named as Goal, Operators, Methods, and Selection (GOMS) which provide a series of methods to evaluate the usability\textsuperscript{26}. Quality in Use Integrated Measurement (QUIM) is a generic comprehensive model used to measure the actual use of the software and identify the problems for refinements\textsuperscript{27}. A Goal Question Metric model (GQM) is used for usability evaluation perspective in multiple areas especially mobile phone applications\textsuperscript{28,29}.

There are many other models and all of them have their own limitations\textsuperscript{30}. Upon close review of the previous studies, the most common factors of usability for this research work are Efficiency, Effectiveness, Satisfaction, Visibility, and Safety. These factors cover almost all the areas of usability for m-Health applications.

#### Usability and m-Health technology

Mobile Health, also known as m-Health, a system being used to provide medicine and services through mobile devices used for improving patients’ lives and their
It also provides an automatic bridge of communication between doctors and patients. Researchers explore m-Health applications for diverse health conditions such as hypertension, insomnia, obesity, Parkinson’s disease. Usability becomes a vital factor in the adoption of these applications. Many researchers explore usability for such applications. Many researchers explore usability for such applications like Xu, in his research proposed usability metrics framework with limited set of usability factors (-I.e; Simplicity, Naturalness, Consistency, Feedback, Effective, Efficient, Cognitive Overload) for m-Health applications. The author only focuses on the operationalized metrics which can only be measured by the software. The author stresses the need of more comprehensive (both operationalized and non-operationalized) usability metrics for m-Health applications.

Another author Estrin focuses on usability for open m-Health architecture of chronic diseases. The author stresses the need of more easy, effective and comprehensive m-health applications. He suggested to expanding the spectrum of current m-health applications architecture by incorporating comprehensive usability guidelines with the broad involvement of patient, physicians, families, diseases. The author also suggested incorporating usability iteratively throughout the development process of such applications.

Liang quantitatively studied the usefulness of hypertension mobile applications for chinese industry. This study is subjective and limited in diseases wise. In the end the author stresses for the exploration of the other usability factors like security, privacy and protection for such applications.

Mirkovic, heuristically evaluated mobile applications for cancer patients. The basic objective of his study was to find the functionality issues of such applications. Seven patients sample sizes were taken for testing in a controlled environment. The author suggested expanding the study by incorporating more usability factors like visibility, usefulness and acceptability.

Zhang in his research have proposed a unified framework of Electronic Health Records (HER) systems usability. The framework was task based which defined, measured and evaluated the usability objectively. Useful, usable and satisfying were taken as basic usability factors for measurement in the framework. The author focuses on the expansion of the representations analysis so that more accurate and visible information will be displayed which is considered to be a basic component for such systems.

Diamantidis investigated the mobile health medication inquiry system for chronic kidney diseases (CDK). The author remotely tested the usability of the e-health (CDK) application and suggested that additional studies are needed for improving the patient safety of such applications.

Another author Eiring has done a comparative usability study on bipolar disorder patients. In this study he highlights the formative and summative effects of such systems on patient’s health. In the implication and future section the author discusses the limitation of the current systems and recommended to deeply investigate the effects of such applications in more comprehensive way.

Fiks evaluated the usability, acceptability, and clinical impact of a tele dermatology mobile applications by linking families with pediatric dermatologists. In the end, the author analyzed that convenience, ease of use, speed of response, and utility have a great impact on the acceptance and rejection of such applications.

Few authors like Stoll, Holzinger and Chan evaluated the usability of m-health applications for mental and elderly patients. The authors presented their findings in the form of metrics and guidelines respectively. The authors suggested integrating usability in the design and development process of such applications.

As describes in literature, a number of researchers explored usability for a different type of m-Health applications, different platforms, and each application shows a number of usability problems. As a result, further studies are needed to investigate usability in a more comprehensive way so that these apps meet their intended user’s needs before being used as health intervention. The technological advancements of mobile devices with a limited capacity of interfacing, battery life, and size also added a new dimension to be explored by the researchers. The effective usability evaluation helps the designers to improve the productivity of such application with low cost and limited time. The usability of any product is determined with user-computer interaction and usability evaluation is considered to be a method...
through which the designers identify the usability problems of their product. There are numerous ways to evaluate the usability either through factors or methods but a limited work is found which systematically evaluate usability through metrics with the combination of evaluations methods and factors for m-Health applications. In this paper, we presented a quality assessment model based on usability metrics for m-Health applications user interfaces. The detailed diagram and description of the model is given in the proposed model section.

RESEARCH METHODOLOGY

This section presents the material and methods used for the development of the proposed model towards its validation. For the development of the proposed model, an extensive literature review has been done. In which multiple usability models, guideline, frameworks, standards (either generic or specific to the research topic) has been explored. Majority of these models uses GQM based approach for the construction of their models. These models are generic and structured in terms of hierarchy28,29,43,44.

Usability Study

For the validation of the proposed model, we use usability testing technique combined with the statistics based analysis. The usability testing is done in a controlled environment in which a measurement instrument in the form of the questionnaire is developed. The instrument is developed on the basis of metrics defined in the proposed model. Two m-Health android based smartphone applications are selected for testing the reliability and effectiveness of the proposed model. The reliability of the instrument is also done statistically by using Cronbach’s Alpha. In usability testing, a total of 10 participants was involved. The demographic detail of the participants is given in Table 1

Table 1: Sample Set demographic details

<table>
<thead>
<tr>
<th>Attributes</th>
<th>Values</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Participants</td>
<td>10</td>
</tr>
<tr>
<td>Gender</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>5</td>
</tr>
<tr>
<td>Female</td>
<td>6</td>
</tr>
<tr>
<td>Qualification</td>
<td>Minimum High School</td>
</tr>
<tr>
<td>Mobile Usage Experience</td>
<td>&gt;3 years</td>
</tr>
<tr>
<td>m-Health app usage</td>
<td></td>
</tr>
<tr>
<td>Experience</td>
<td>Novice 6</td>
</tr>
<tr>
<td></td>
<td>Expert 5</td>
</tr>
</tbody>
</table>

Two applications named Lose Weight App and Diabetes: M apps were taken as case studies. The no of participants was selected according to the Nigel Bevan theory who suggested that 8 to 10 participants are required to make reliable estimates of usability problems associated with user interface45. Each participant is instructed to use the apps according to the required tasks and fill the post-test questionnaire using five (5) points Likert Scale that ranges from 1 to 5, where 1 means low and 5 means high. The moderator records the participant’s comments and other details. Both m-Health applications were given to the participants on a rotational basis. The tasks detail for each application is given in Table 2. The final outcomes of the tests are given in the result section.

<table>
<thead>
<tr>
<th>Table 2: Task List</th>
<th>Tasks</th>
</tr>
</thead>
<tbody>
<tr>
<td>Application Title</td>
<td>Tasks</td>
</tr>
<tr>
<td>Lose Weight In 21 Days:</td>
<td>Open the application</td>
</tr>
<tr>
<td>Home Fitness Workouts</td>
<td>Select the training level after reading the help literature</td>
</tr>
<tr>
<td></td>
<td>Set the plan of exercise for 2 days</td>
</tr>
<tr>
<td></td>
<td>Choose exercises as back, legs and arms 2(for each)exercises</td>
</tr>
<tr>
<td></td>
<td>See the steps of each exercise</td>
</tr>
<tr>
<td>Diabetes: M</td>
<td>After login check your sugar level</td>
</tr>
<tr>
<td></td>
<td>Check your glucose history through graphs</td>
</tr>
<tr>
<td></td>
<td>Check how much food you will take</td>
</tr>
<tr>
<td></td>
<td>Send report to the Doctor</td>
</tr>
<tr>
<td></td>
<td>Check the log book of current day details</td>
</tr>
</tbody>
</table>

Proposed Model.

For our proposed model we used the GQM based approach where we classify our proposed model into goals, factors and metrics layers which is hierarchically structured. The goals and factor were selected from the literature and specific to the domain (m-Health). The metrics were developed according to the factors and further classified as subjective or objective. These metrics helped in measuring the objective and subjective usability of m-Health applications. The diagrammatical representation of the proposed model is given in Fig 1. The detailed description of each goal is given below.

Efficiency: It is considered to be a capacity of the system to produce the desired results as a response of the resources invested in it17,18,39,46. The resources were in
the form of tasks and human effort. For m-Health applications, it is explored as the resources expended in the relation of task, memory, time and navigability.

**Effectiveness:** It is defined as the system capacity to react consistently, flexible, error-free with better memorability and comprehensively in features (utility)\(^{17,18,39,43}\).

**Satisfaction:** This goal is subjective in nature and it represents the system capacity to provide attractiveness, productivity, helpfulness and readability with maximum user control\(^{17,18,39,46,47}\). In m-Health applications user need control so that they take their medical decisions in a more controlled and knowledgeable manner.

**Visibility:** The m-Health applications use many medical images for the learning, diagnosis and other purposes. The low visibility does not provide a sufficient freedom to the user through which they can lean or take their medical decisions. This goal is considered to be as a system capacity to visualize and present the medical contents (i.e. medical literature, images, and diagrams) in a more meaningful way\(^{11,36,48,49}\).

**Safety:** Many m-Health applications are keeping their patients records\(^{5,11,36,50,51}\). Patients are very much concerned about the privacy and safety of their medical records. This goal has a capacity to keep records safe and secure with fault tolerance facility\(^{11,36,48,52}\).

**RESULTS AND DISCUSSION**

The quality models based on usability metrics cannot be validated directly\(^{53}\). The validation of these models is based on the no of usability problems identified by these models\(^{54}\). In this paper we conducted a usability study on two m-Health applications, the mean based results of both cases studies are represented graphically in Fig 2 and 3. Fig 4 represents the comparative analysis of both applications.

An independent-samples t-test was conducted to compare both applications. The results show that there is a significant difference in the means of both applications where \(t = 0.267\) which is > 0.05. The result positively supports the argument that m-health applications are aligned with the proposed model and proposed models have a capability to compare and target multiple usability problems for different nature of applications. The reliability of the instrument is also tested using Cronbach’s Alpha.
technique. The Cronbach's Alpha value of both applications is .28 and .26 which is > 0.6000 indicated that the items in the instrument are reliable.

The overall analysis of the results shows that the proposed model is not only useful for comparing different applications but it also identifies the usability problem areas for further improvements.

**CONCLUSION**

In this paper, we proposed a quality assessment model which offers a comprehensive structure to assess the quality in terms of usability of m-Health applications. A comprehensive review of various usability models, frameworks, and attributes for m-health applications have been done to develop usability goals and metrics. The model described the usability issues specifically needed for m-Health applications and integrate those issues in the form of usability characteristics to design the user interfaces. The proposed metrics become a useful tool for the evaluators and usability designers to obtain the quantitative and qualitative data for evaluating the usability of m-Health applications.

**Future Aspiration**

In the future, this model can be enhanced by adding more usability goals based on future technological trends and demands of the user. The metrics provided by the proposed model can also be validated for usability case studies with a large group of participants in a real-time environment. These metrics can also be used to develop more evaluation instruments such as checklists, heuristics, and guidelines.

**REFERENCES**


44. Santosso SS. Penguaran Aspek Efisiensi Usability Aplikasi Mobile SatNav Menggunakan Metode Mobile Goal Question Metrics (mGQM) Pada Device iPhone 4S dan Garmin Nüvi 1460i. Jurnal Ilmiah Universitas Bakrie. 2015; 3(01).


**Appendix-I**

**Figure 1: Proposed Quality Assessment Model for m-Health Applications**

- **Efficiency**
  - Task Efficiency
  - Task Efficacy
  - Time Behavior
  - Memory Load
  - Learnability
  - Navigability

- **Effectiveness**
  - Consistency
  - Flexibility
  - Error
  - Memorability
  - Utility

- **Usability**
  - Helpful
  - Attractive
  - User Control
  - Productivity
  - Readability

- **Satisfaction**
  - Visibility
  - Presentation

- **Safety**
  - Record Security
  - Fault Tolerance

- **Usability**
  - Estimate no of Clicks acquired to Complete a Task
  - Estimate Total No of Tasks Completed
  - Estimate the number of tasks with associated tasks completed
  - Time taken to complete total tasks
  - Estimate subjective workload (e.g., estimate Stress, time taken to complete individual tasks)
  - Estimate the degree of system to engage user
  - Estimate easiness of the user to learn and operate the system
  - Estimate the user proficiency to move within the application
  - Estimate the uniformity of different elements of user interface
  - Estimate the meaningful representations of different user interface elements
  - Estimate the number of alternative options provided by the system to complete one task
  - Estimate the total errors done by the user while accomplishing the task
  - Estimate the System error prevention capability for the user
  - Estimate system help options (customization and shortcut keys option)
  - Estimate system help features provide an adequate guidance to the user
  - Estimate the system readability features to provide an adequate guidance to the user
  - Measure how the users easily remember the performed tasks
  - Estimate the System error prevention capability for the user
  - Measure how easy for the user to understand the health literature
  - Measure the graphical representations attractiveness to user while performing the task
  - Estimate how much user feel confident in their ability to perform tasks
  - Estimate the fruitfulness of the information provided by the system
  - Measure how easy for the user to read the health literature
  - Measure how easy for the user to take its medical decision
  - Estimate how much user personal medical information is correctly protected
  - Estimate the system capacity to recover
  - Estimate the system capacity to handle unexpected errors (e.g., timing data input)