SUDAN UNIVERSITY OF SCIENCE AND TECHNOLOGY

COLLEGE OF GRADUATED STUDIES

COLLEGE OF COMPUTER SCIENCE AND INFORMATION TECHNOLOGY

A Proposed Usability Evaluation Checklist and Threshold for Interactive Systems

مقترح قائمة تدقيق و قيمة مرجعية لتقييم إستخدامية الأنظمة التفاعلية

A Thesis Submitted in Partial Fulfillment of the Requirements of M.Sc. in Computer Science

PREPARED BY:-
Ebtehal Hashim EL-khalifa

SUPERVISED BY:-
Dr. Nisreen Beshir Osman

April 2018

الآيـــــــــــــة

(وَلَوْلَا فَضْلُ اللَّهِ عَلَيْكَ وَرَحْمَتُهُ لَهَمَّتْ طَائِفَةٌ مِنْهُمْ أَنْ يُضِلُّوكَ وَمَا يُضِلُّونَ إِلَّا أَنْفُسَهُمْ ۖ وَمَا يَضُرُّونَكَ مِنْ شَيْءٍ ۚ وَأَنْزَلَ اللَّهُ عَلَيْكَ الْكِتَابَ وَالْحِكْمَةَ وَعَلَّمَكَ مَا لَمْ تَكُنْ تَعْلَمُ ۚ وَكَانَ فَضْلُ اللَّهِ عَلَيْكَ عَظِيمًا)

سورة النســـاء الآيـة 113

Dedication

Special dedication to my family members especially to my beloved father and mother (Hashim EL-Khalifa and Ehssan Hassan) who always give me encouragement in my life, my study and to finish my research.

To my Supervisor

Dr. Nisreen Beshir

To all my classmates

And all my friends

Thank you for your support

Thank you for everything that you have given me during my studies and the knowledge that we shared together.

Abstract

Designing for maximum usability is the goal of interactive systems design, since users want interactive products to be easy to learn, effective, efficient, safe, memorable and over all; satisfying to use. Achieving this requires the product to be evaluated, but the process of evaluation can be difficult, because different evaluation methodologies require some restriction “e.g. experienced evaluators” and can be more time-consuming. In addition, developers will not be able to know how their developing product compares to widely known, highly used products since the usability of an artifact is defined by the context in which that artifact is used. So there is a need to develop appropriate usability evaluation measures that allow system developers to make more informed evaluation with their own systems quickly as well as comprehensively regardless of their context. Hence, a usability evaluation checklist has been developed. The proposed checklist has been designed based on design rules in views of User Interface design elements. A hierarchical structure of UI design elements and usability principles were developed and then utilized to develop the checklist. Most usable products were included in this study to be evaluated using the developed checklist to provide the developers with a new threshold that will aid them in determining where their tested products fall within a distribution of highly used products. These products are one of the top sites on the web including YouTube, Facebook and Wikipedia. Moreover, products that are best in class and that are used in every-day life as Microsoft Office. Finally, the heading product in specific domain such as W3Schools, which is the number one online education source for beginner in Web developing. The result of the evaluation found that the agreed assessment result for the newly developed product must exceed the threshold of the proposed checklist to be considered as a usable product.

المستخلص

يعتبر التصميم لقابلية الاستخدام القصوى هو الهدف الاساسي من تصميم الأنظمة التفاعلية، حيث يريد المستخدمين ان تتميز هذه الانظمة بسهولة التعلم، الكفاءة والفعالية، الآمان، وعلى وجه العموم ان تكون هذه الانظمة مرضية الاستخدام. لتحقيق ذلك، يجب ان يتم تقييم النظام التفاعلي المطور. ولكن من المحتمل ان تكون عملية التقييم صعبة او مجهدة، لان منهجيات التقييم المختلفة قد تتطلب بعض القيود مثال لذلك 'مقيمون خبراء' وقد تستهلك ايضا زمناً طويلا. بالإضافة الى ذلك، من الصعب على مطوري الأنظمة تحديد موقع منتجهم مقارنة بالمنتجات المعروفة و عالية الاستخدام. وذلك لان قابلية الإستخدام لأي منتج تحدد بحسب السياق لذلك المنتج. لذلك دعت الحاجة الى تطوير منهجية مناسبة لتقييم قابلية الاستخدام للأنظمة التفاعلية كي يستطيع مطورو هذه الأنظمة أن يقيموها بصورة سريعة وشاملة بغض النظر عن سياقها. تم إقتراح قائمة تدقيق لتقييم قابلية الأستخدام. صممت قائمة التصميم المقترحة بناءاً على قواعد التصميم وإستناداً الى عناصر التصميم لواجهات المستخدم. وقد تم تطوير هيكل هرمي لعناصر تصميم واجهة المستخدم ومبادئ قابلية الاستخدام ومن ثم استخدامها لوضع قائمة التدقيق. في هذه الدراسة ليتم تقييمها بإستخدام قائمة التدقيق المطورة تم تضمين معظم المنتجات عالية الاستخدام لتوفيرمقياس لمطوري النظم ليساعدهم في تحديد وضع منتجهم بالنسبة للمنتجات عالية هذه المنتجات هي واحدة من أكثر المواقع وجهة بالنسبة للمستخدمين على شبكة .الإستخدام الانترنت مثل 'يوتيوب، فيسبوك و ويكيبيديا'. بالاضافة لذلك المنتجات التي تعتبر الافضل في صنفها مثل حزمة 'مايكروسوفت'. واخيرا المنتجات التي تعتبر المتصدرة في مجال '، حيث يعتبر هذا الموقع مصدر التعليم الأول على الانترنت بالنسبة W3Schoolsمعين ً لمطوري الويب المبتدئين. وجدت النتائج أن المنتج المطور بعد تقييمه بقائمة التدقيق المقترحة، يجب ان يتجاوز القيمة المرجعية المقترحة ليتم اعتباره منتج عالي الإستخدام.

List of Contents

List of Table

List of Figures

List of Abbreviations

UI	User Interface
HCI	Human Computer Interaction
ODG	Office Design Group
UX	User Experience
UT	Usability Testing
WebCT	Web Course Tools

1. Chapter One

Introduction

1.1. Background

Usability is the extent to which users can use a computer system to achieve a specific goal in an effective and efficient way while promoting a feeling of satisfaction in a given context of use (ISO, 1998). This characteristic is important for the success of interactive systems since the users may want interactive systems to be easy to learn, effective, efficient, safe, memorable and over all; satisfying to use. Achieving this, requires the product to be evaluated, but the process of evaluation can be difficult, because different usability evaluation methodologies require some restrictions "e.g. experienced evaluators" and can be more time-consuming. In addition, developers will not be able to know how their developing product compares to widely known, highly used products since the usability of an artifact is defined by the context in which that artifact is used. Hence, the focus of this study is to identify an opportunity for improving the traditional heuristic evaluation technique.

1.2. Problem Statement

1- Different usability evaluation methodologies require some restrictions "e.g. experienced evaluators, task-based" and can be more time-consuming (Holzinger, 2005).

2- Because the usability of an artifact is defined by the context in which that artifact is used (Brooke, 1996). Hence, when developing a new product, developers will not be able to know how their product compares to widely known, highly used products 'e.g. Microsoft Word', which results in difficulty to determine if the usability of a new product is comparable and/or has exceeded an agreed-upon relative threshold.

1.3. Research Significance

There is an existing growing demand to explore appropriate evaluation methodology that evaluates the usability quickly as well as comprehensively with minimum cost. In addition, there is a need for broad general measures that can be used to compare usability across a range of contexts.

1.4. Objectives

1- Develop a usability evaluation checklist that makes the process of evaluation of developing interactive system quickly as well as comprehensively with minimum cost.

2- Establish threshold for widely-known, highly used products to allow systems developers to make more informed comparisons with their own products regardless to their context.

1.5. Scope of the Study

1.5.1. Design Rules Included Scope

Important Design Rules of the design process were selected. First, the principles. Principles are abstract design rules, with high generality (applied to many design situations or focused on specific application situation) and low authority (whether or not a rule must be followed or whether it is just suggested). Then, golden rules and heuristics. Useful checklist or summary of the essence of design advice are provided. There are many sets of heuristics, but the best used are Nielsen's ten heuristics, Shneiderman's eight golden rules and Norman's seven principles (Dix, 2009).

1.5.2. Samples Measured Scope

Five of common, widely known products were chosen to characterize their level of usability (YouTube, Facebook, Wikipedia, Microsoft Office and W3Schools).

1.6. Thesis Organization

The study has structured as the following:

Chapter 2 Provide Background and Literature review of the thesis

Chapter 3 Contains the Research Methodology.

Chapter 4 Shows the Evaluation and Results Discussion.

Finally, Chapter 5 Concludes the study.

2. Chapter Two

Literature Review

Introduction

This chapter has divided into two section; literature review and previous studies related to this study.

2.1. The Concept of Usability

Usability is a measure of interface quality that refers to the effectiveness, efficiency and users satisfaction, so they can perform tasks with a tool. Historically, the concept of usability has been defined in various ways (Dillon, 2001):

- Semantics: here, usability is equated to terms such as 'ease of use' or 'user-friendliness', without formal definition of the properties of the structure.
- Features: in this case, usability is equated to the presence or absence of certain features in the user interface such as Windows, Menus or Pointing devices.
- Operations: which is defined in terms of performance and affective levels clear by users for certain task and environmental scenarios.

The first type of definition was useless for design purposes since it offers neither useful guidance for designers nor perspective for evaluators. The feature-based approach rests on an assumption that usability is an inherent part of the application; this assumption is false since one could always consider a combination of users, with certain task demands, in a specific environment, for whom a given set of features would be suboptimal. Hence, most human factors professionals now employ an operational definition in their work, which explicitly places usability at the level of the interaction between users and the artifact. This takes it beyond the typical features-based definitions common in

the field. Furthermore, in setting criteria for assessing usability, this approach better supports the evaluation of any tool and the subsequent interpretation of the test results. Usability therefore refers not to a set of interface features, but to a context-dependent measure of HCI (Dillon, 2001).

2.2. Usability evaluation

Usability evaluation assesses the extent to which an interactive product is easy and pleasant to use. HCI researchers and Interaction Design professionals have developed evaluation methods that determine whether or not an interactive product is usable. Where a system is usable, usability evaluation methods also determine the extent of its usability, through the use of robust, objective and reliable metrics.

Evaluation methods and metrics are thoroughly documented in the HCI research and practitioner literature. People wishing to develop expertise in usability measurement and evaluation can read about these methods, learn how to apply them, and become proficient in determining whether or not an interactive product is usable, and if so, to what extent (The Interaction Design Foundation, 2017)

2.2.1. Usability Methods

There are multiple methods of evaluating usability depending on available resources, evaluator experience, ability and preference, and the stage of development of the tool under review. Here, following distinctions between evaluation methods (Dillon, 2001):

- User-based: where a sample of the intended users try to use the application.
- Expert-based: where an Human Computer Interaction (HCI) or usability expert makes an assessment of the application
- Model-based: where an HCI expert employs formal methods to predict one or more criteria of user performance

2.2.2. Expert-Based

Expert-based methods refers to any form of usability evaluation which involves an HCI expert examining the application and estimating its likely usability for a given user population. In such cases, the evaluation lies in the interpretation and judgment of the evaluator, which mean that the users are not employed. There is considerable interest in this form of evaluation since it can produce results faster and presumably cheaper than user-based tests.

Two common expert-based usability evaluation methods are Heuristic evaluation (e.g., (Nielsen, 1994)), and Cognitive Walkthrough (Wharton et al, 1994). Both methods aim to provide evaluators with a structured method for examining and reporting problems with an interface (Dillon, 2001).

2.2.2.1. Heuristic Evaluation

"A heuristic is a guideline or general principle or rule of thumb that can guide a design decision or be used to critique a decision that has already been made" (Dillon, 2001).

The Heuristic method provides a simple list of design guidelines, which the evaluator uses to examine the interface screen by screen and while following a typical path through a given task. The evaluator reports violations of the guidelines as likely user problems. Heuristic methods are based on design guidelines and ultimately reflect the expert's judgment of how well the interface conforms to good design practice (Dillon, 2001).

2.2.3. Design Rules

They are rules a designer can follow in order to increase the usability of the eventual software product. There are number of different types of design rules. Principles are abstract design rules, with high generality. Standards are specific design rules, limited in application, whereas Guidelines tend to be more general in application.

Design rules are mechanisms for restricting the space of design options, preventing a designer from pursuing design options that would be likely to lead to an unusable system. Hence, design rules would be most effective if they could be adopted in the earliest stages of the software life cycle, such as in requirements specification and architectural design, when the space of possible designs is still very large.

Principles are derived from knowledge of the psychological, computational and sociological aspects of the problem domains and they are largely independent of the technology; depending to a much greater extent on a deeper understanding of the human element in the interaction. They can therefore be applied widely but are not so useful for specific design advice (Dix et al., 2009).

2.2.3.1. Principles that Support Usability

Principles are divided into three main categories (Dix et al., 2009):

- Learnability: the ease with which new users can begin effective interaction and achieve maximal performance. Table 3.2 shows principles affecting learnability.
- Flexibility: the multiplicity of ways in which the user and system exchange information. Table 3.2 shows principles affecting Flexibility.
- Robustness: the level of support provided to the user in determining successful achievement and assessment of goals. Table 3.2 shows principles affecting Robustness.

2.2.3.2. Golden Rules and Heuristics

To aid the evaluators in discovering usability problems, useful checklist or summary of the essence of design advice are provided. There are many sets of heuristics, but the widely used are Nielsen's ten heuristics (Nielsen, 1994), Shneiderman's eight golden rules (Shneiderman, 1998) and Norman's seven principles (Norman, 1998) (Dix et al., 2009).

Table 2.1 Principles affecting learnability

Principle	Definition
Predictability	Support for the user to determine the effect of future action based on past interaction history
Synthesizability	Support for the user to assess the effect of past operations on the current state
Familiarity	The extent to which a user's knowledge and experience in other real-world or computer based domains can be applied when interacting with a new system
Generalizability	Support for the user to extend knowledge of specific interaction within and across applications to other similar situations
Consistency	Likeness in input–output behavior arising from similar situations or similar task objectives

Table 2.2 Principles affecting Flexibility

Principle	Definition
Dialog initiative	Allowing the user freedom from artificial constraints on the input dialog imposed by the system
Multi-threading	Ability of the system to support user interaction pertaining to more than one task at a time
Task migratability	The ability to pass control for the execution of a given task so that it becomes either internalized by the user or the system or shared between them
Substitutivity	Allowing equivalent values of input and output to be arbitrarily substituted for each other
Customizability	Modifiability of the user interface by the user or the system

Table 2.3 Principles affecting Robustness

Principle	Definition
Observability	Ability of the user to evaluate the internal state of the system from its perceivable representation
Recoverability	Ability of the user to take corrective action once an error has been recognized
Responsiveness	How the user perceives the rate of communication with the system
Task conformance	The degree to which the system services support all of the tasks the user wishes to perform and in the way that the user understands them

2.3. Style Guide

Style guides can be classified as platform (or language) guide (focus on rules for presentation elements, including visual design elements; page or screen layouts and common items; and the correct usage for standard controls), or general design guides which are looking at the overall structure of a site and how the user navigates through it as an important design element. They may also include a primary focus on the process for creating a usable interface design. Or corporate style guides such as hybrid of platform guides and design guides for a specific application to help the product meet business and usability requirements (Quesenbery, 2001).

2.4. Most Usable Products

The products to be evaluated were chosen from among the most usable products. These products were chosen according to several criteria:

1- Product that is the top sites on the web destination as YouTube, Facebook and Wikipedia (Alexa, 2017).

2- Product which is best in class and that are used in every-day life as Microsoft Office (Microsoft Office, 2017).
3- The heading product in specific domain as W3Schools, which is the number one online education source for beginning Web developers (Eu.wiley, 2017).

2.4.1. Microsoft Office

Microsoft Office is an office suite of applications, servers, and services developed by Microsoft. Bill Gates first announced it on 1 August 1988, at COMDEX in Las Vegas. Initially a marketing term for a bundled set of applications, the first version of Office contained Microsoft Word, Microsoft Excel, and Microsoft PowerPoint. In Microsoft office, there is a group called Office Design Group (ODG). ODG made up of both User Experience Designers and User Experience Researchers. Their job is to represent you, the end-user of their software products. They collaborate with product teams within Office to identify user needs and create compelling experiences. By understanding who you are and how you work, they can build better software. ODG includes User experience (UX) Designers who work to create compelling software. UX design defines how software looks and behaves. They are deeply interested in the interaction models that affect how software is perceived, learned and used (Microsoft Office, 2017).

One component of the Office suite was selected; Microsoft Word which is a graphical word processing application.

2.4.2. YouTube

YouTube is an American video-sharing website headquartered in San Bruno, California. The service was created by three former PayPal employees — Chad Hurley, Steve Chen, and Jawed Karim — in February 2005. Google bought the site in November 2006; YouTube now operates as one of Google's subsidiaries and now YouTube site is considered as number two on the web destination (Alexa, 2017).

2.4.3. Facebook

Facebook is a social networking website and service where users can post comments, share photographs and links to news or other interesting content on the Web, play games, chat live, and even stream live video. Facebook began in February of 2004 as a school-based social network at Harvard University. It was created by Mark Zuckerberg along with Edward Saverin (Lifewire, 2017) . Facebook site is considered as number three on the web destination (Alexa, 2017).

2.4.4. W3Schools

W3Schools is a web developer's site, which is providing tutorials and references on web development languages such as HTML, CSS, JavaScript, PHP, SQL, W3.CSS, and Bootstrap, covering most aspects of web programming. Refsnes Data, a Norwegian software development and consulting company (w3schools, 2017) originally created W3Schools in 1998.

W3Schools is the top Google search result for instruction on HTML, CSS, and other key Web technologies (Eu.wiley, 2017).

2.4.5. Wikipedia

is a multilingual, web-based, free-content encyclopedia project supported by the Wikimedia Foundation and based on a model of openly editable content.it considers the largest and most popular reference web-site in the word. Wikipedia is also unique as this encyclopedia is written by everyone and can be read by anyone (Google Books, 2017). 2.4.5. Wikipedia site considers number five on the web destination (Alexa, 2017).

2.5. Related Work

Several studies in usability have been carried out, either the study in usability assessment or in usability's tool and technique assessment. This section mentions some of work that has been done in the field of usability.

2.5.1. Heuristic Evaluation on Mobile Interfaces: A New Checklist:

(Gómez et al., 2014) presented a compilation of heuristic evaluation checklists readapted to mobile interfaces. They started their work by reusing heuristics from desktop heuristics evaluation checklists, which is allowed because “heuristic checklists change very slowly, since they derive from human behavior, not technology” (Budiu & Nielsen, 2011) In fact, in the final proposal of this work, the amount of reused heuristics from the literature is 69% of the total proposed sub heuristics. The rest are best practices and recommendations for mobile interfaces not initially conceived as part of a usability tool. The result is a comprehensive checklist, which is experimentally evaluated as a design tool. This experimental evaluation involved two software engineers without any specific knowledge about usability, a group of ten users who compared the usability of a first prototype designed without (Gómez et al., 2014) proposed heuristics, and a second one after applying their proposed checklist. The results of this experiment show the usefulness of their proposed checklist for avoiding usability gaps even with non-trained developers.

2.5.2. Academic Library Website Design Principles: Development of a Checklist:

(Raward, 2001) showed 'as user' that acceptance and usability are major issues in the design of library websites; hence, this work suggests that the design will be most successful when a user-centered design model is included in the development and implementation of academic library web pages. This work suggests that the use of a HCI usability checklist (developed from principles derived from the HCI literature), such as the one designed by Keevil (Raward, 2001), if used during the design process of a website, it will ensure that best design practice principles are supported. To further assist academic library website, designers have modified version of the Keevil checklist, which has been designed specifically to address the issues involved in an academic library website, this work has been suggested as the tool best suited to ensure user-centered best practice principles.

2.5.3. Heuristic Evaluation for Games: Usability Principles for Video Game Design:

(Pinelle, 2008) introduced a new set of heuristics that can be used to carry out usability inspections of video games. The heuristics were developed to help in identifying usability problems in both early and functional game prototypes. They have developed the heuristics by analyzing PC game reviews from a popular gaming website, and the review set covered 108 different games and included 18 from each of 6 major game genres. They have analyzed the reviews and identified twelve common classes of usability problems seen in games. They have developed ten usability heuristics based on the problem categories, which are describing how common game usability problems can be avoided. The heuristics were created by translating the problems into principles that provide guidance on how they can be avoided. For example, the problem category ' unpredictable/inconsistent response to user's actions' became the heuristic ' provide consistent responses to user's actions '.

2.5.4. A Usability Checklist for the Usability Evaluation of Mobile Phone User Interface:

(Ji et al., 2010) developed 21 usability principles that are crucial in the mobile phone UI design. This study aims to develop a task-based usability checklist based on heuristic evaluations in views of mobile phone user interface (UI) practitioners. A hierarchical structure of UI design elements and usability principles related to mobile phones were developed and then utilized to develop the checklist. In the proposed usability checklist, there exists a promising methodological benefit to enhance the effectiveness and efficiency of heuristic evaluation; hence, comparative experiments were conducted on the usability checklist and User Testing (UT). The result of comparative experiments on both usability checklist and UT revealed that about 90% of usability problems identified by UT were covered by their proposed usability checklist. According to the literature (Law & Hvannberg, 2002), heuristic evaluation reports typically do not predict 30 to 50% of usability problems

found by UT. Given the result in (Ji et al., 2010) study, it can be asserted that their proposed checklist evaluation might improve the traditional heuristic evaluation technique.

2.5.5. An Adaptable Usability Heuristic Checklist for Online Courses:

(Dringus et al., 2005) conducted a heuristic evaluation of WebCT (Course Tools) from the user perspective, including a faculty and student perspective. The authors as online professors conducted a heuristic evaluation of WebCT in which over 100 usability problems were located within an hour. The authors made an independent list of the usability problems found, connected the problems to Nielsen and Mack's usability heuristics, and then aggregated the two problem lists into one comprehensive list. From the aggregated problem list, the authors identified 13 heuristic categories and specific heuristics that match each category. The 13 heuristic categories are: Visibility, Functionality, Aesthetics, Feedback and Help, Error Prevention, Memorability, Course Management, Interactivity, Flexibility, Consistency, Efficiency, Reducing Redundancy, and Accessibility. The result of the usability evaluation is an adaptable usability heuristic checklist in its draft stage.

3. Chapter Three

Research Methodology

This chapter presents the methodology adopted in this study; a hierarchical structure of UI design elements and usability principle/design rules were developed and then utilized to develop the checklist. The procedure includes eliciting principles from the range of usability principles to be match with specific UI design element according to specific steps to develop the usability checklist at the end as shown in Figure 3.1. This chapter describe each processes achieved in development procedure.

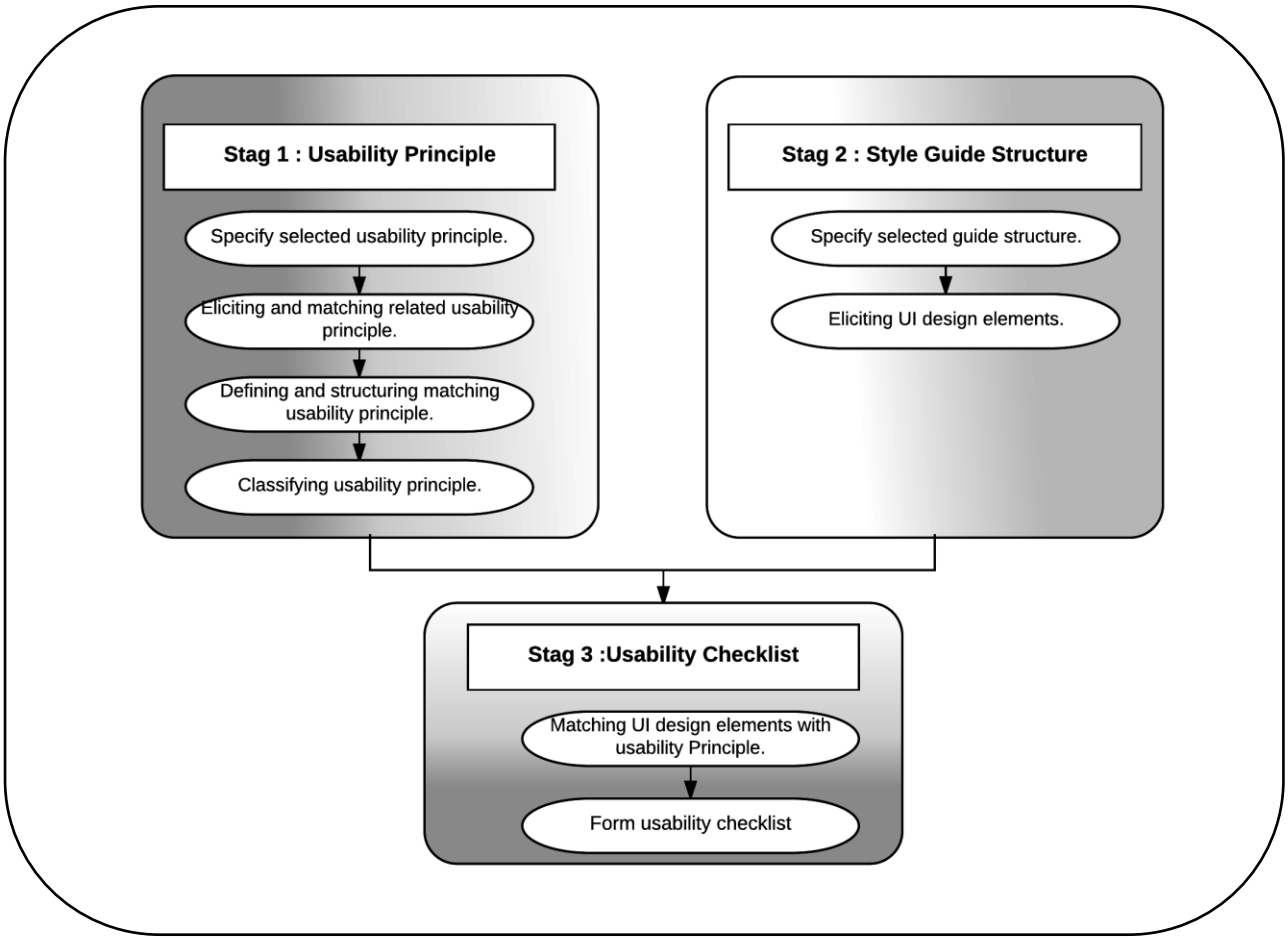

Figure 3.1: Development procedure for usability checklist.

3.1 Stage1: Usability Principles

3.1.1. Specify Selected Usability Principles

Important usability principles of the design process have been arranged: the principles. Principles are abstract design rules, with high generality either applied to many design situations or focused on specific application situation and low authority (whether or not a rule must be followed or whether it is just suggested). Then, golden rules and heuristics which are a useful checklist or summary of the essence of design advice provided to aid the evaluators in discovering usability problems. There are many sets of heuristics, but the best used are Nielsen's ten heuristics, Shneiderman's eight golden rules and Norman's seven principles. For instance (Shneiderman, 1998) and (Nielsen, 1994) focused on several principles, such as visibility of system status, consistency, and freedom. Table 3.1 fully describes the selected Usability Principles and their sources.

3.1.2. Eliciting and Matching Related Usability Principles

Arranged usability principles are carefully elicited and regarded as important usability principles that must be considered in the software UI design process. (Dix et al., 2009) have selected all the arranged usability principles except Generalizability principle, since we find that Applying generalization to situations in which the user wants to apply knowledge that helps achieve one particular goal to another situation where the goal is in some way similar. Here comes consistency, likeness in input–output behavior arising from similar situations or similar task objectives, so generalizability can be seen as a form of consistency accordingly, generalizability was integrated into consistency.

Then the eliciting usability principles were matched based on the similarity in the purpose of the principles. Table 3.2 shows the matching usability principles.

Table 3.1: Selected Usability Principles

Principle	Definition
(Dix et al., 2009)	***Learnability:*** Predictability; Synthesizability; Familiarity; Generalizability; Consistency. ***Flexibility:*** Dialogue initiative; Multithreading; Task migratability; Substitutivity; Customizability. ***Robustness:*** Observability; Recoverability; Responsiveness; Task conformance.
(Shneiderman ,1998)	Strive for Consistency؛Cater to Universal Usability؛ Offer Informative feedback؛ Design Dialogs to yield closure؛ Prevent Errors؛ Permit easy reversal of actions؛ Support internal locus of control؛ Reduce short term memory load.
(Nielson, 1994)	Visibility of system status؛ Match between system and the real world؛ User control and freedom؛ Consistency and Standards؛ Help users recognize, diagnose, and recover from errors؛ Error Prevention؛ Recognize rather than recall؛ Flexibility and efficiency of use؛ Aesthetic and minimalist design؛ Help and documentation.
(Norman, 1998)	Use both knowledge in world and knowledge in the head؛ Simplify task structures؛ Make things visible؛ Get the mapping right; Exploit the power of constraints؛ Design for Error؛ When all else fails – Standardize.

3.1.3. Defining and Structuring Matching Usability Principles

Matching usability principles are carefully defined and structured, For example, Familiarity by (Dix et al., 2009) and match between system and the real world by (Nielsen, 1994) agree on how prior knowledge applies to new system. Hence principles were matching to yield one principle contain all considerations for involved principles called Familiarity. The same term was

Table 3.2: The Matching Usability Principles.

Dix et al.		Shneiderman	Nielson	Norman
Learnability	**Predictability**	- Support internal locus of control		
	Synthesizability	- Offer Informative feedback	- Visibility of system status	- Simplify task structures - Make things visible
	Familiarity		- Match between system and the real world	- Use both knowledge in world & knowledge in the head
	Generalizability			
	Consistency	- Strive for Consistency - Offer Informative feedback	- Consistency and Standards	- When all else fails – Standardize - Exploit the power of constraints
Flexibility	-	- Enable frequent users to use shortcuts	- Flexibility and efficiency of use	
	Dialogue initiative	- Support internal locus of control		- Simplify task structures
	Multithreading			
	Task migratability	- Support internal locus of control		
	Substitutivity			- Simplify task structures
	Customizability	- Support internal locus of		

		control		
Robustness	-	- Prevent Errors.	- Errors Prevention.	
	Observability	- Design Dialogs to yield closure	- Visibility of system status	
Robustness	***Recoverability***	- Permit easy reversal of actions.	- User control and freedom - Help users recognize, diagnose, and recover from errors	- Design for Error
	Responsiveness		-Visibility of system status	
	Task conformance			
-	-	- Reduce short term memory load	- Recognize rather than recall	- Simplify task structures
-	-		-Aesthetic and minimalist design;	
-	-		- Help and documentation.	

kept for best expression. Another example, Responsiveness by (Dix et al., 2009) and Visibility of system status by (Nielsen, 1994) agreed on how the user perceives the rate of communication with the system, so they have been matched to yield one principle called Responsiveness. The same term was kept also for best expression. Furthermore, with Reduce short term memory load by (Shneiderman, 1998), Recognize rather than recall by (Nielsen, 1994), and Simplify task structures by (Norman, 1998), all of them are concerned with the concept of recognition, so they were put together and were called; Recognition,

and so on. Some principles have no match with others principle, for example Help and documentation By (Nielsen, 1994), so they have been considered as a separated principle and was called User Support. Table 3.3 shows The Definition and Structure of Matching Design Rules.

Table 3.3: The Definition and Structure of Matching Design Rules

Design Rules		Definition
Learnability	***Predictability***	- Determine the effect of Operations on the system -Determine which operations can be performed (operation visibility). - Avoid non-causality.
	Synthesizability	- Assess the consequences of previous interactions which change some aspect of the internal state (substantial &immediate vs. Modest &eventual feedback). - Indication of how long and how much system operation, which takes some time, is complete. - Provide more information about the task, what it does and how this was achieved.
	Familiarity	- First impression of the system (Metaphor, Affordances). - Culture (language, user-oriented terms, information appear in natural and logical order).
	Generalizability	- Generalizability can be seen as a form of consistency.
	Consistency	- Consistency in interface element, input expressions or output Responses and feedback.
Flexibility	***Enable shortcuts***	- Provide abbreviations, special key sequences.
	Dialogue initiative	- Maximize the user's ability to pre-empt the system (user imitate dialogue) and minimize the system's ability to pre-empt the user (system imitate dialogue).
	Multithreading	- Ability of system to support user interaction for several tasks at a time.

	Task Migratability	- Possibility for the user or system to pass the control of a task over to the other.
	Substitutivity	- Allowing equivalent values of input and output to be substituted for each other
	Customizability	- Modifiability of the user interface by the user (adaptability) or system (adaptivity)
Robustness	***Error Prevention***	- UIs that make it hard to make errors by allowing only suitable option in certain situation.
	Observability	- Ability of the user to evaluate the internal state of the system from its perceivable representation and knows when they have completed a task.
	Recoverability	- Ability of the user to correct a recognized error. - Leave the unwanted state without having to go through an extended dialog. - Reachability states: Forward and backward recoveries (redo / undo). - Commensurate effort states: Require more effort by the user to do actions. - Error messages expressed in plain language, indicate the problem, and suggest a solution.
	Responsiveness	- How the user perceives the rate of communication with the system.
	Task conformance	- Degree to which system services support all of the user's tasks.
Recognition	- Instructions for use of the system should be visible or easily retrievable whenever appropriate. - Keeping displays simple. - Provide mental aids to help the user keep track of stages in a more complex task.	
Structured Design	- Keep interface simple and organized.	
Minimalist Design	- Dialogs should not contain information that is irrelevant or rarely needed.	
User support	- Printed/On-line./Training./Help ‘desk’.	

3.1.4. Classifying Usability Principles

After the usability principles were selected, they were classified to make matching between UI design elements and usability principles more abstract. According to classification made by (Gómez et al., 2014), four classifications of them representing the selected principles:

1. Cognition Support relates to cognitive aspects of user.
2. Information Support relates to characteristics of UI display and information.
3. Interaction Support relates to the interaction between user and UI.
4. User Support relates to the degree of intervention of user. Table 3.4 shows the classification of usability principle.

Table 3.4: Classification of Usability Principle

Classification	Usability Principle	
Cognition support	Learnability	Predictability
		Synthesizability
		Familiarity
		Generalizability
		Consistency
User support	Flexibility	Enable shortcuts
		Dialogue initiative
		Multithreading
		Task Migratability
		Substitutivity
		Customizability
Interaction support	Robustness	Error Prevention
		Observability
		Recoverability
		Responsiveness
		Task conformance

Information support	Recognition
Information support	Minimalist Design
User support	User Support

3.2. Stage2: Style Guide Structure

3.2.1. Specify Selected Guide Structure

(Quesenbery, 2001) style guide has been selected because it presents important design element and also include a primary focus on the process for creating a usable interface design.

3.2.2. Eliciting UI Design Elements from Style Guide Structure

From the selected style guide structure (Quesenbery, 2001)., UI design elements that have direct effect in usability of representational content were elicited, then to make the evaluation process simpler, UI elements were split to more concrete image so the process of matching with the classified usability principle will be more accurate. Table 3.4 shows the hierarchical structure of UI design elements.

3.3. Stage3: Form Usability Checklist

3.3.1. Matching UI Design Elements with Usability Principles

Since the target usability principle and specific UI element to be evaluated have reached, hence, we can go to the final step to develop the usability checklist. Each classified usability principle, which has a direct effect on the selected and elicited UI element; has matched. Table 3.6 shows the matched usability principle with the UI Design elements.

3.3.2. Form Usability Checklist

The matched usability principle and elicited UI element have been utilized to develop a checklist; Table 3.7 shows the proposed usability evaluation checklist.

Table 3.5: Structure of UI Design Elements

First level	Second level	Third level
Architecture or Structure	Menus or Control Bars	▪ Menu bar ▪ Menu ▪ Context menu
	Start Page	▪ Home Page (or Desktop or Main Menu)
Page or Window Layout	Page Structure	▪ Size of page margins. ▪ Size and position of images and figures. ▪ Deciding on the number and size of columns and gutters (gaps between columns). ▪ Placement of intentional whitespace. ▪ Use of special effects. ▪ Boxouts and sidebars. ▪ Page headers and page footers.
	Window Types	▪ Overlapped Windows ▪ Pop-up Windows ▪ Child Windows ▪ Layered Windows ▪ Message-Only Windows
Interactions	Keyboard Shortcuts	▪ General shortcuts ▪ Navigation shortcuts ▪ Text editing and formatting shortcuts ▪ Web browsers shortcuts
	Interaction task and feedback type	▪ Input ▪ confirmation ▪ backward/cancel/termination. ▪ load.
User Assistance	Messages	▪ Status Bar Messages ▪ Error Messages
	User Assistance	▪ Key Metaphors ▪ Tooltips ▪ Embedded Help
Visual Design	Logos	▪ Balanced design ▪ Color ▪ Typography ▪ Different design
	Typography	▪ Fonts ▪ Size

		▪ Alignment ▪ Color ▪ Spacing ▪ Labels and Prompts
	Colors	▪ Differentiate items ▪ Create depth ▪ Add emphasis ▪ Organize information ▪ Meaning per color
	Icons Styles	▪ Recognizability ▪ Consistently ▪ Standing-out ▪ Metaphors and Affordance ▪ Uniquely
Multimedia Design	Static Elements	▪ Text ▪ Image
	Dynamic Elements	▪ Video ▪ Audio ▪ Animation

Table 3.6: Matched Related Usability Principle and UI Component

***UI* Component**	***Principle***			
	Cognition Support	*Information Support*	*Interaction Support*	*User Support*
Architecture or Structure	- Consistency	- Recognition - Structured - Minimalist Design	- Error Prevention	- Customizability
Page or Window Layout	- Consistency	- Structured -Minimalist Design		
Interactions	- Predictability - Synthesizability - Consistency		- Error Prevention - Observability - Recoverability - Responsiveness -Task conformance	- Enable shortcuts - Dialogue initiative - Multithreading - Task Migratability - Substitutivity

User Assistance and Text	- Consistency - Familiarity			- User support
Visual Design	- Consistency - Familiarity	- Recognition - Structured -Minimalist Design		
Multimedia Design	- Consistency	- Recognition		

Table 3.6: The Proposed Usability Evaluation Checklist

First Level Element	*Second Level Element*	*Evaluation Items*	*Evaluation*					*Note*
			Never applied 0	*Poorly applied 1*	*Neutral 2*	*Mostly applied 3*	*Fully applied 4*	
Architecture Or Structure	Menus or Control Bars	- Is the Menu bar consistent in every window? - Are The menus easy to recognize? - Are the menus preserve only suitable options in certain situation? - Can users change the menu, as they desire?						
	Start Page	- Is Start Page simple and organized? - Is Start Page containing only relevant information?						
Page or Window Layout	Page Structure	- Is Size of page margins consistent in certain window and every window? - Is the number and size of columns and gutters consistent in certain window and every window? - Is the Placement of intentional whitespace consistent in certain window and every window? - Is the Use of special effects consistent in certain window and every window? - Are the Boxouts and sidebars consistent in certain window and every window? - Page headers and page footers consistent in every						

		window? - Is the Design elements in page structure are simple and organized? Is every Dialog in certain window containing only information that is relevant or frequently needed?		
	Window Types	- Is there consistency between windows in a certain type of windows?		
Interactions	Keyboard Shortcuts	- Are there good shortcuts that are Meaningfully related to longer procedures that are provided?		
	Interaction task and feedback type	- Are input expressions that can be performed are determined? - Is there consistency in input expressions? -Are there preserved to only suitable input expressions in certain situation? - Is there possibility for equivalent values of input to be substituted for each other? - Are the consequences of previous interactions, which change some aspect of the internal state, can be assess in case of conformation of the task? - Are the unwanted state easily recognized and corrected (or cancelled/ backward) by the user? - Is unwanted state can be leaved /cancelled without having to go through an extended dialog? - Is there possibility for Reachability states		

		(Forward/ redo and backward/ undo) to be done in case of unwanted state? - Are the non-retrieved actions done by the user Requiring more effort to be done? - When user makes action/change, is the notifications/feedbacks are observable and come immediately requiring no further interaction initiated by the user? - For long delays to complete an action/change, is there detailed progress indication are Provided? -Is there consistency between notifications/feedbacks for similar action/change? - Is the user entirely free to initiate any action toward the system unless for system control reasons which it may be necessary to prohibit the user from the 'freedom' to do potentially serious damage? - Are there tasks that perfectly suited to automation so the overhead on the user can be decreased? - Are all of tasks of the user interest in certain system domain are supported?		
User Assistance	Messages	- Are there a status bar Message to display information about the current state of its window? - Is there consistency between error messages (color/shapes used/place)? - Is an error message easy to recognize?		

	User Assistance	- Is there metaphor/affordance used in system? - Is there user documentation and support (Printed/On-line/T raining/Help 'desk')?		
Visual Design	Logos	- Is there balancing in "weight" of the graphics, colors, and size equal on each side? - Is the logo also looks good in black and white, grayscale, and two colors? - If there exist Typography in logo, is the most commonly used fonts avoided? - If there exist Typography in logo, is the font legible when scaled down? - If there exist Typography in logo, are the fonts used are tow or less? - Is the logo simple to be easy to recognize? - Is the logo has distinct style that makes it stand out from the competition?		
	Typography	- Is there consistency in using font/size/alignment/color/spacing on typography? - Is the used font familiar to users? - Is the color used in typography easy to recognize? - Are the used Languages, user-oriented? - Are the used terms, user-oriented? - Are the spacing and alignment have used in a way that make structured interface? - Are the font types used are four or less? - Are the Sizes used are four or less? - Are the colors used are four or less?		

	Colors	- Are there colors used to Differentiate items? - Are there colors used to Create depth? - Are there colors used to Add emphasis? - Are there colors used to Organize information? - Meaning per color? - Are the colors used are five or less? - Is the user culture about the colors concerned?		
	Icons Styles	- Are the Icons Styles standing out making it easy to recognize? - Are the Icons Styles have designed according to any metaphor/affordance? - Is each icon unique in the system?		
Multimedia Design	Static Elements	- Is the image standing out making it easy to recognize? - Is there consistency between images (size/place)?		
	Dynamic Elements	- Is the video standing out making it easy to recognize? Is there consistency between videos (size/place)? - Is the audio standing out making it easy to recognize? - Is the animation standing out making it easy to recognize? Is there consistency between animations (size/place)?		

4. Chapter Four

Evaluation and Results Discussion

This chapter presents the evaluation process of the proposed checklist and the discussion of the evaluation results.

4.1 Evaluation Experiment for Usable Products

The products have been evaluated evaluating using the proposed usability checklist to provide the developers with a new threshold that will aid them in determining where their tested products fall within a distribution of highly used products.

4.1.1 Subject/Participants Profile

The participants have been involved in this experiment were three (two male and one female) of master students, college of Computer Science and Information Technology in Sudan University of Science and Technology. They have knowledge in the evaluated products, moreover in human computer interaction concept.

4.1.2 Instructions to Participants

Participants were given the proposed checklists and were asked to assess the selected usable systems by going through the system and evaluate the aspects of the system according to the evaluation element, and to give a weight for each evaluation element depending on which the degree an aspect of the evaluated system applied the certain evaluation item. Figure 4.1 shows the ranging of evaluation elements weight.

0	1	2	3	4
Never applied	Poorly applied	Neutral	Mostly applied	Fully applied

Figure 4.1 Ranging weight for evaluation item

4.1.3 Method

The candidates had assessed three product, Microsoft Word, YouTube, and W3Schools.Table .4.1shows the result of the assessment. Every evaluation element of the proposed checklist evaluates specific aspects of the product, for example one of the aspect of Microsoft Office World can be measure by the function ‘replace’ , where the proper evaluation element will be ‘When user makes action/change, is the notifications/feedbacks are observable and come immediately requiring no further interaction initiated by the user?’. In Figure 4.2 and Figure 4.3, it was found that the observation of the replaced word (‘num’ replaced with ‘No.’) was not immediately, as we go through the text to check out the result of the replacement resulting in possibility of mistakes (e.g. ‘number’ replaced to ‘No.er’) that couldn’t be detected immediately, hence the rate of the evaluation element will decrees. Another example, in YouTube, the evaluation element ‘Are the non-retrieved actions done by the user requiring more effort to be done?’ can be measured with the process of deleting video from your channel, as shown in Figure 4.4 and Figure 4.5, the process done with an extended dialog, hence the rate of the evaluation elements will increase, and so on.

The assessment results for each product have been acquired from each participant, then, the averages of assessment result have been conducted for each product. the results shows that Facebook has the highest average of all 91.51%, coming next Microsoft word scores 87.46%, third comes YouTube 83.77%, then W3Schools scores 72.42%, and the final one is Wikipedia with score 67.73%. Table 4.1 and Figure 4.6 shows the Scores for the each of the five products.

Once we have accrued the averages that have been obtained from the evaluation of most usable product, we can draw scale that can be used as an agreed-upon threshold for the proposed evaluation checklist by getting the **Standard Deviation** of the evaluated selected usable product. The equation to get the Standard Deviation is (Hoel&P.G., 1960):

Table 4.1: The Result of Usable Product Assessment

Product	Assessment results %			Average %
	Participant 1	Participant 2	Participant 3	
Microsoft Word	89.03	86.06	87.288	87.46
YouTube	81.56	85.47	84.27	83.77
Facebook	89.77	91.54	93.23	91.51
W3Schools	69.50	73.66	74.09	72.42
Wikipedia	70	68.45	64.73	67.73

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2}.$$

Where
σ is the standard deviation
$\mathbf{x_i}$ is an individual value
μ is the mean/expected value
N is the total number of values

Hence,

$$\mu = (87.46+83.77+91.51+72.42+67.73) / 5 = 80.58$$

$$\sigma = \sqrt{}\,[(87.46 - 80.58)^2 + (83.77- 80.58)^2 + (91.51- 80.58)^2 + (72.42- 80.58)^2 + (67.73- 80.58)^2]/5 = 9.04$$

9.04 is the standard deviation of the averages for the selected usable product, which mean the scale that can be used as an agreed-upon threshold is in the range of 71.18 to 89.62.

4.1.4 Results Discussion

The results show that the agreed assessment result found by the proposed usability evaluation checklist for the newly developed product should be in the proposed scale, which is from 71.18% to 89.62%, this scale has been

developed by evaluating products which consider as usable products all over the word. For more flexibility in evaluation process; it could be said that the threshold of the proposed checklist is 71%. Therefore, the developed product must exceed this threshold to be considered as a usable product.

But the noticed that some products have low assessment result, e.g. Wikipedia as shown in Table .4.1. This means that some system are not usable because the usability principles were considered during design process but could be usable because they are unique in providing specific services to the users 'one-of-class', hence the user well surely use the product and get used to it.

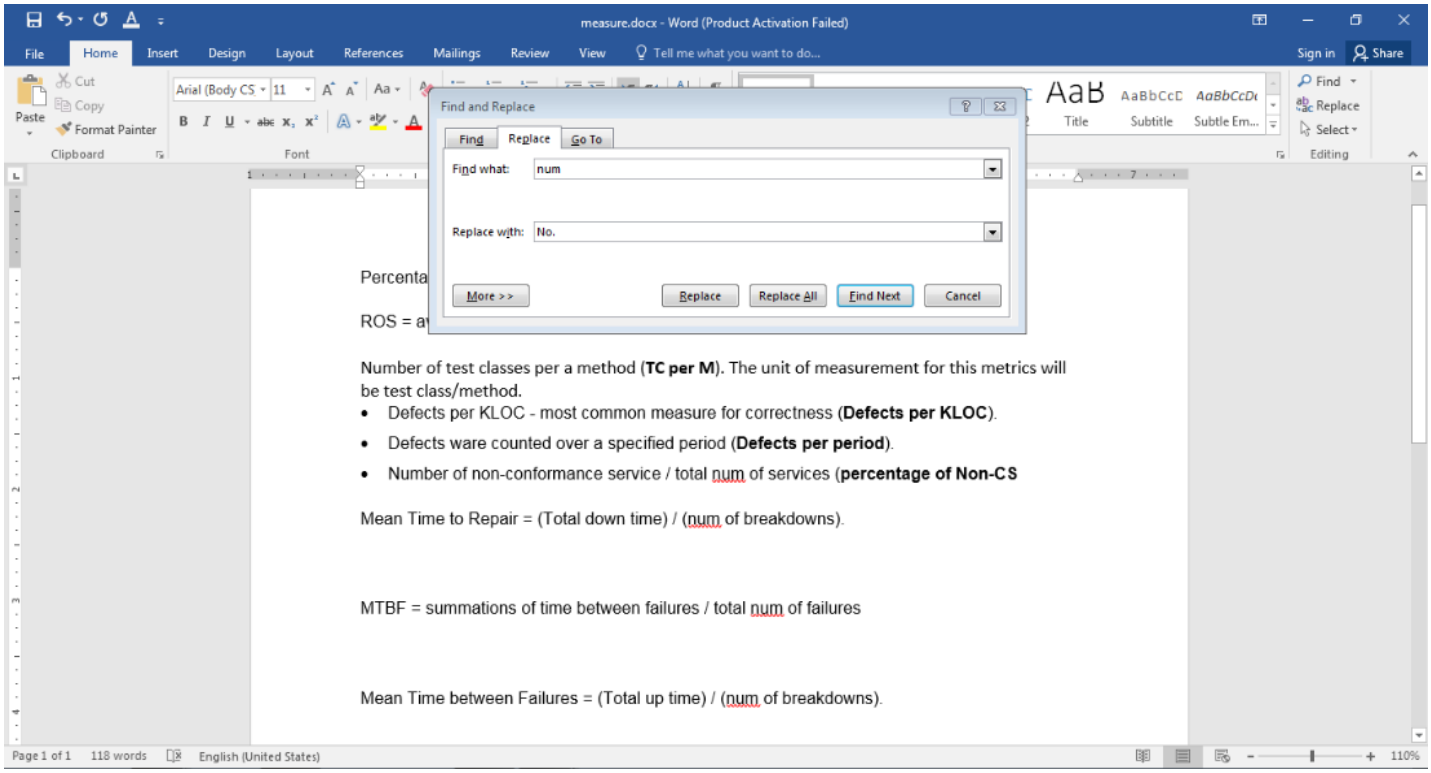

Figure 4.2 A screen of evaluating Microsoft office world 2013

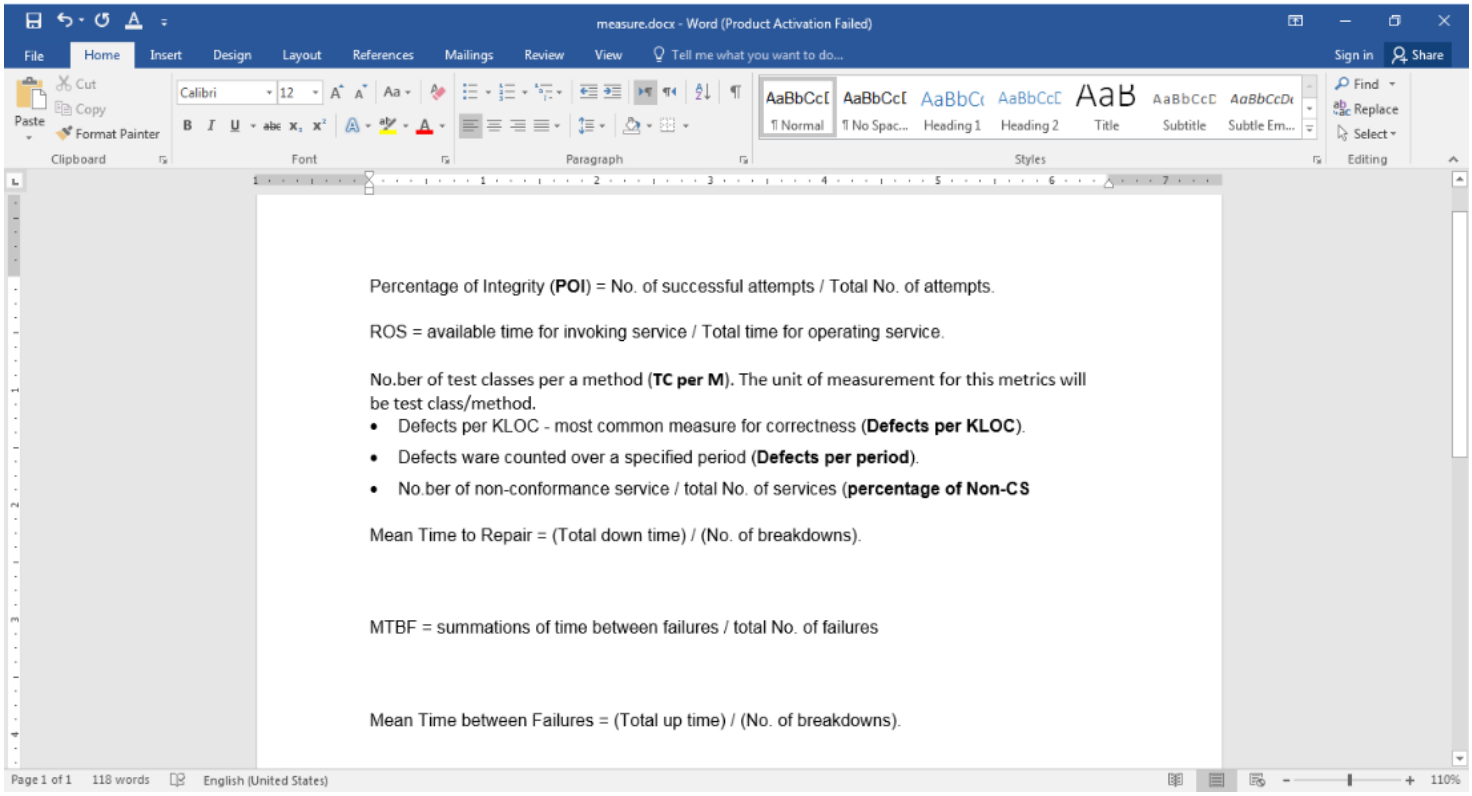

Figure 4.3 A screen of evaluating Microsoft office world 2013

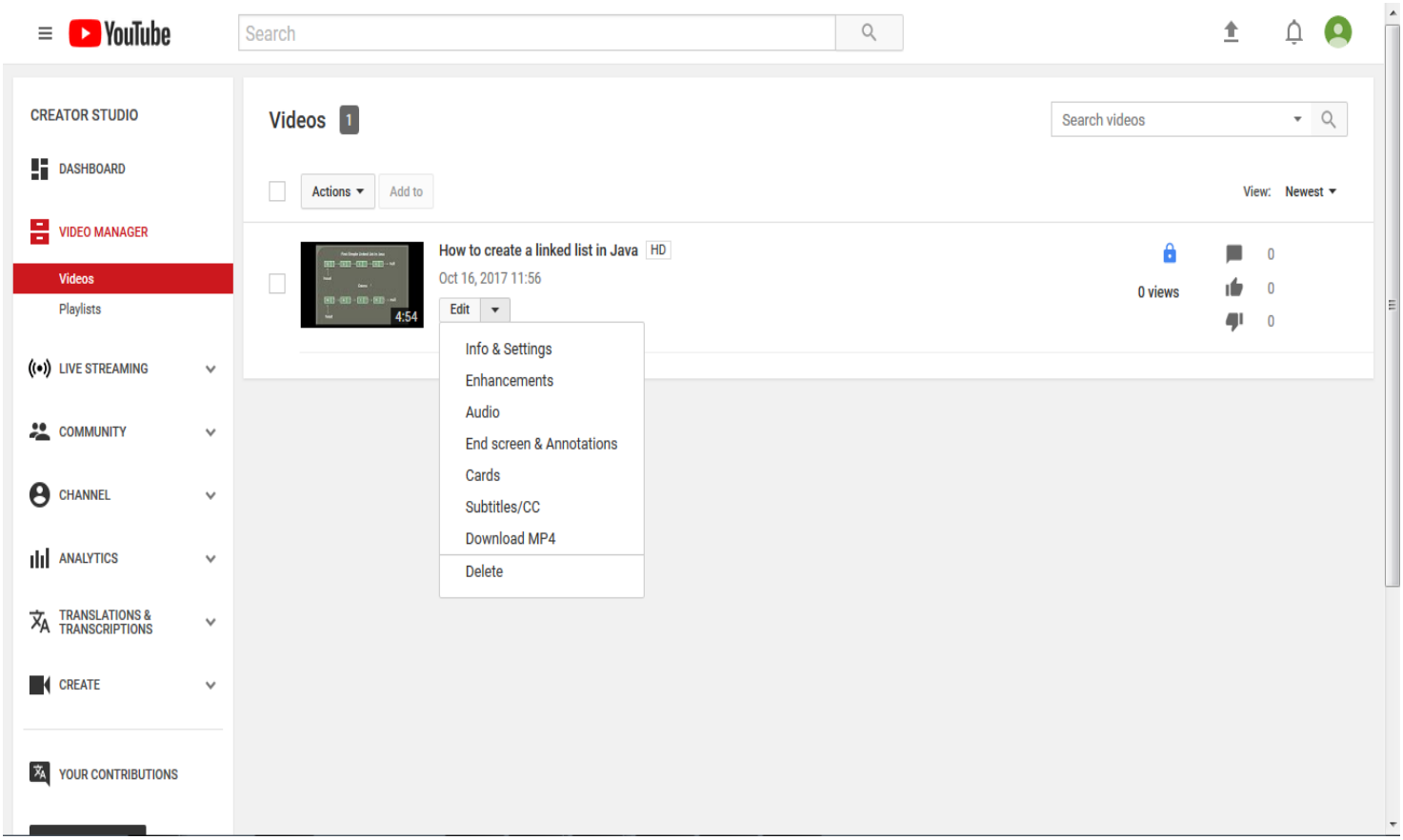

Figure 4.4 **A** screen of evaluating YouTube site

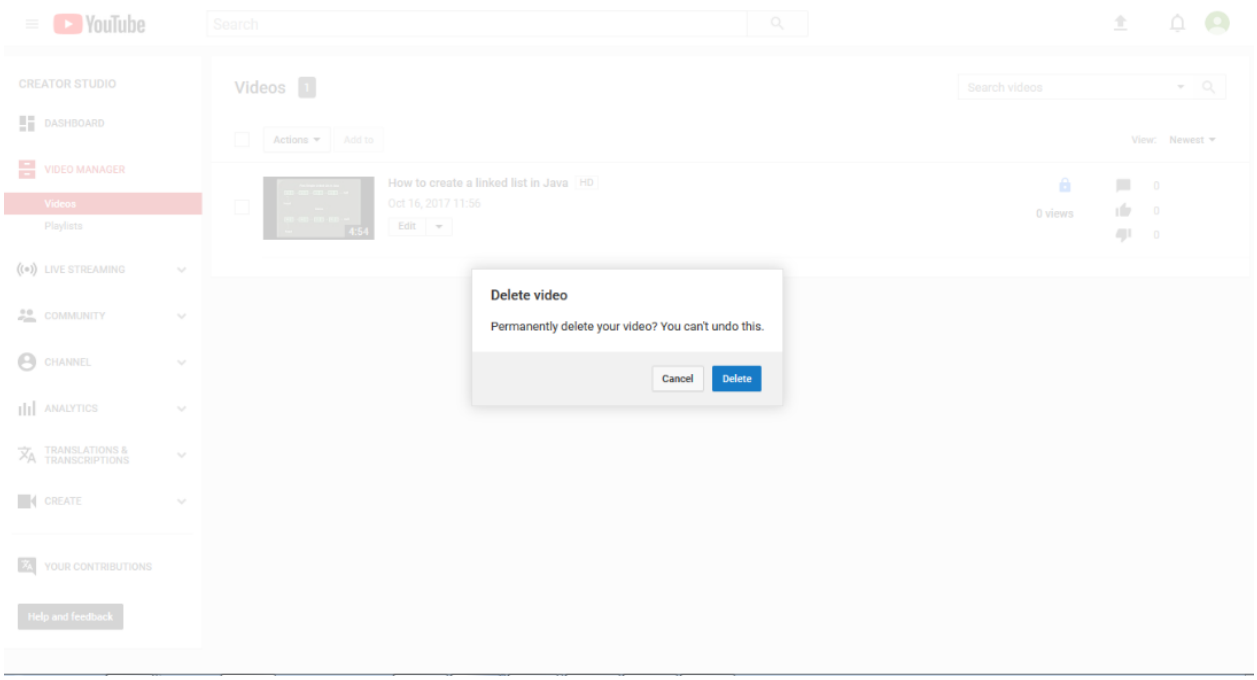

Figure 4.5 **A** screen of evaluating YouTube site

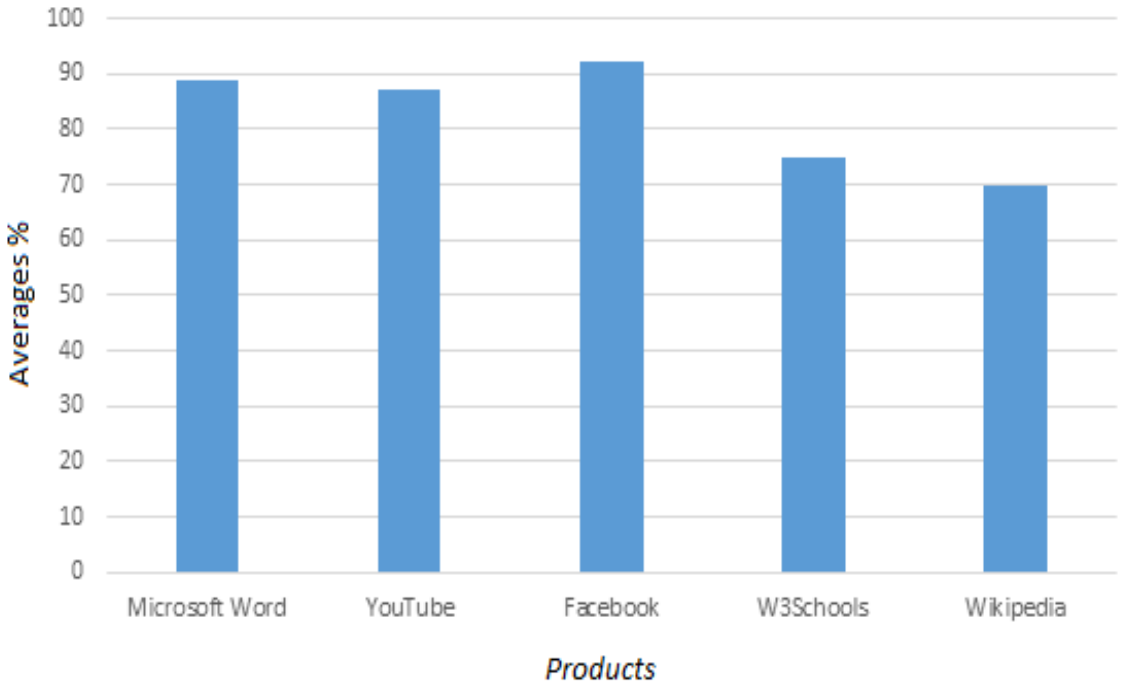

Figure 4.6 the Averages of the Usable Product Assessment

5. Chapter Five

Conclusion and Recommendations

5.1 Conclusion

Usability is the most important characteristic in interactive systems, since users want interactive products to be satisfying to use. Achieving this requires the product to be evaluated, but the process of evaluation can be difficult since different evaluation methodologies require some restriction “e.g. experienced evaluators”and can be more time-consuming, in addition, mostly they are context-based. This study proposes a usability evaluation checklist that evaluates the usability of developing product quickly as well as comprehensively regardless to their context, in other word, the focus of this study is to identify an opportunity for improving the traditional evaluation technique. The proposed checklist was designed based on Design rules in views of UI design elements. A hierarchical structure of UI design elements and usability principles were developed and then utilized to develop the checklist. Most usable product were included in this study to be evaluated using the developed checklist to demonstrate the efficiency of it, and also to provide the developers with a new threshold that will aid them in determining where their tested products fall within a distribution of highly used products. The result of evaluation found out the agreed assessment result for the newly developed product must exceed the threshold of the proposed checklist, which is 71%, to be considered as a usable product.

5.2 Recommendations

Its recommended that to demonstrate the practical effectiveness of the proposed checklist, comparative experiments should be conducted on the usability checklist and UT which is a way to see how easy to use something is by testing it with real users (Users are asked to complete tasks, typically while

they are being observed by a researcher, to see where they encounter problems and experience confusion. If more people encounter similar problems, recommendations will be made to overcome these usability issues.) (Experience UX, 2017). To compare if the majority of usability problems found by usability testing can be discovered by the proposed checklist, and even more an additional problems or not. Moreover, to test if the proposed checklist can discover with regard to the various problems of specific UI elements found by it the serious problems of interaction occurring, that can discover by the UT.

6. References

Alexa, 2017, Alexa Top 500 Global Sites. [online] Available at: https://www.alexa.com/topsites [Accessed 13 Sep. 2017].

Brooke, J., 1996. SUS-A quick and dirty usability scale. Usability evaluation in industry, 189(194), pp.4-7.

Dillon, A., 2001. The evaluation of software usability. In Encyclopedia of human factors and ergonomics. London: Taylor and Francis.

Dix, A., 2009. Human-computer interaction. In Encyclopedia of database systems (pp. 1327-1331). Springer US.

Dringus, L.P. and Cohen, M.S., 2005, October. An adaptable usability heuristic checklist for online courses. In Frontiers in Education, 2005. FIE'05. Proceedings 35th Annual Conference (pp. T2H-6). IEEE.

Eu.wiley, 2017, Wiley: Learn HTML and CSS with w3Schools - W3Schools. [online] Available at: http://eu.wiley.com/WileyCDA/WileyTitle/productCd-0470611952.html [Accessed 13 Sep. 2017].

Experience UX, 2017, What is usability testing? | Experience UX. [online] Available at: http://www.experienceux.co.uk/faqs/what-is-usability-testing/ [Accessed 19 Oct. 2017].

ISO, W., 1998. 9241-11. Ergonomic requirements for office work with visual display terminals (VDTs). The international organization for standardization, 45.

Google Books, 2017, How Wikipedia Works. [online] Available at: https://books.google.com/books?id=lHdi1CEPLb4C&pg=PP1&hl=en#v=onepage&q&f=false [Accessed 1 Oct. 2017].

Hoel, P.G., 1960. Elementary statistics. Elementary statistics.

Holzinger, A., 2005. Usability engineering methods for software developers. Communications of the ACM, 48(1), pp.71-74.

Ji, Y.G., Park, J.H., Lee, C. and Yun, M.H., 2010. A usability checklist for the usability evaluation of mobile phone user interface. International journal of human-computer interaction, 20(3), pp.207-231.

Lifewire, 2017, What Is Facebook? Here's What You Should Know. [online] Available at: https://www.lifewire.com/what-is-facebook-3486391 [Accessed 19 Oct. 2017].

Law, L.C. and Hvannberg, E.T., 2002, October. Complementarity and convergence of heuristic evaluation and usability test: a case study of universal brokerage platform. In Proceedings of the second Nordic conference on Human-computer interaction (pp. 71-80). ACM.

Microsoft Office, 2017, New Ways to Try and Buy Microsoft Office 2010. [online] Available at: http://blogs.technet.com/b/office2010/archive/2009/10/07/new-ways-to-try-and-buy-microsoft-office-2010.aspx [Accessed 8 Jul. 2017].

Nielsen. Heuristic evaluation. In Usability Inspection Methods. John Wiley, New York, 1994.

Norman, the Design of Everyday Things, MIT Press, 1998. Classic exposition of design, including the seven principles.

Pinelle, D., Wong, N. and Stach, T., 2008, April. Heuristic evaluation for games: usability principles for video game design. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (pp. 1453-1462). ACM.

Quesenbery, W., 2001. Building a better style guide. In Proceedings of UPA.

Raward, R., 2001. Academic library website design principles: development of a checklist. Australian Academic & Research Libraries, 32(2), pp.123-136.

R. Budiu and N. Nielsen, Usability of iPad Apps and Websites, Nielsen Norman Group, 2nd edition, 2011.

Shneiderman, Designing the User Interface, 3rd edition, Addison-Wesley, 1998. The source of Eight Golden Rules.

The Interaction Design Foundation, 2017, Usability Evaluation. [online] Available at: https://www.interaction-design.org/literature/book/the-encyclopedia-of-human-computer-interaction-2nd-ed/usability-evaluation [Accessed 3 Jun. 2017].

W3schools.com, 2017 , About W3Schools. [online] Available at: https://www.w3schools.com/about/default.asp [Accessed 19 Oct. 2017].

Yáñez Gómez, R., Cascado Caballero, D. and Sevillano, J.L., 2014. Heuristic evaluation on mobile interfaces: a new checklist. The Scientific World Journal, 2014.