Palestine Polytechnic University
Deanship of Graduate Studies and Scientific Research
Master of informatics

A Study of Usability of Arabic Forms for Smartphones

Submitted by:

Rihabb Ramadan Alsalamen

Thesis submitted in partial fulfillment of requirements of the degree Master of Science in Informatics

August 2017

The undersigned hereby certify that they have read, examined and recommended to the Deanship of Graduate Studies and Scientific Research at Palestine Polytechnic University the approval of a thesis entitled: **A Study of Usability of Arabic Forms for Smartphones**, submitted by Rihabb R. Alsalamin in partial fulfillment of the requirements for the degree of Master in Informatics.

Graduate Advisory Committee:

Dr.Ghassan Shahin (Supervisor), Palestine Polytechnic University.

Signature: _____ Date: _____

Dr. Mahmoud AL-Saheb (Internal committee member), Palestine Polytechnic University.

Signature: _____ Date: _____

Dr. Mamoun Nawahdah (External committee member), Birzeit University.

Signature: _____ Date: _____

Thesis Approved

Dr. Murad Abu Sbeih
Deanship of Graduate Studies and Scintific Research
Palystine Polytichnic University

Signature: _____ Date: _____

DECLARATION

I declare that the Master Thesis entitled "A Study of Usability of Arabic Forms for Smartphones" is my original work, and herby certify that unless stated, all work contained within this thesis is my own independent research and has not been submitted for the award of any other degree at any institution, except where due acknowledgement is made in the text.

Rihabb Ramadan Alsalamen

Signature: _____

Date: _____

STATEMENT OF PERMISSION TO USE

In presenting this thesis in partial fulfillment of the requirements for the master degree in Informatics at Palestine Polytechnic University, I agree that the library shall make it available to borrowers under rules of the library.

Brief quotations from this thesis are allowable without special permission, provided that accurate acknowledgement of the source is made.

Permission for extensive quotation from, reproduction, or publication of this thesis may be granted by my main supervisor, or in his absence, by the Dean of Graduate Studies and Scientific Research when, in the opinion of either, the proposed use of the material is for scholarly purposes.

Any coping or use of the material in this thesis for financial gain shall not be allowed without my written permission.

Rihabb Ramadan Alsalamen

Signature: _____ Date: _____

DEDICATION

To my tutors for their aspiring guidance. To my friends and colleagues who supported me all along. To my family for every single drop of their time, energy, patience, and encouragement. And above all ... to my mother, my saving grace. The one remarkable woman who raised me to be the person who I am today.

I dedicate my work and humble efforts for every person ... a friend or a foe who has ever passed by and has led to create this astonishing moment. To those who made the dream become a reality ... and those who made this very moment possible.

I simply thank you all.

ACKNOWLEDGEMENT

As my dissertation finalization stage come to an end, and as I put my finishing touch on my work, I would like to direct my deep appreciation and sincere gratitude to my supervisor Dr. Ghassan Shahin for his tremendous efforts, for his constructive criticism, and for his endless support during my journey in the research arena.

I also would like to take a moment to reflect on the people who facilitated the entire process. I would like to acknowledge the efforts made by the students that participated in the experiments. I am grateful for their dedication.

Last but not least, I would like to thank my family and friends who have sticked by me tirelessly and helped me during the data gathering stage.

الملخص

لقد أقر الكثير من الباحثين أن فحص استخدامية المنتج بأشكاله المختلفة تعد عملية غاية في الأهمية، سواء كان هذا المنتج جهاز او موقع الكتروني او حتى صفحة الكترونية منفردة التي قد تحوي نماذج ادخال ذات أغراض مختلفة. تطرق الكثير منهم لدراسة قابلية استخدام النماذج بكثرة خلال العقدين السابقين. من الجدير ذكره ان الأعمال السابقة في هذا المجال كانت مرتبطة بالنماذج الانجليزية والتي يتم استخدامها بواسطة الحاسوب. بعد انبثاق التطور الهائل في مفهوم الحاسوب وظهور ما يسمى بالأجهزة الذكية، التي يمكنها القيام بالكثير من مهام الحاسوب والتي تتراوح من العمليات الحسابية المعقدة الى العمليات البسيطة كتصفح الانترنت، الأمر الذي أدى الى الحاجة الى عملية فحص قابلية استخدام النماذج الخاصة بهذه الأجهزة، التي تتمتع بخصائص عدة، كصغر حجمها وامكانية التنقل خلال استخامها. بعد الاطلاع على الأبحاث السابقة، وجدنا بأن هنالك القليل من الأبحاث التي تحدثت في هذا المجال، كما ان أغلب الأبحاث تتمحور حول نماذج اللغة الانجليزية. هذه الأمور دفعتنا الى القيام بالبحث في النماذج العربية التي يتم استعمالها بواسطة الهواتف الذكية. نحن نسعى في هذه الدراسة الى تحقيق مجموعة من الأهداف. أولاً: دراسة تأثير حجم الشاشة على معدل الأخطاء التي يرتكبها المستخدم خلال عملية ادخال بياناته بواسطة النماذج العربية. ثانياً: دراسة وتحديد أفضل توقيت ومكان لظهور رسالة الخطأ. وأخيراً: دراسة تأثير تظليل حقل الادخال في لحظة حدوث الخطأ، بالاضافة الى ظهور رسالة الخطأ كما هو معتاد باللون الأحمر. اثبتت نتائج هذا البحث ان مستخدمي الشاشات الصغيرة يرتكبون عدد أكبر من الأخطاء من مستخدمي الشاشات الكبيرة. كما أظهرت النتائج ان أفضل توقيت لرسالة الخطأ هو ظهورها مباشرة بعد الانتهاء من تعبئة الحقل وليس النموذج كاملاً. تمكن مستخدمي نماذج الفئة الأولى من انهاء مهامهم في فترة زمنية اقل من الفئة الثانية وبعدد اخطاء اقل. علاوة على ذلك فقد فضل معظم المستخدمين اظهار رسائل الخطأ بعد الانتهاء مباشرة من تعبئة الخطأ. اما بالنسبة لأفضل موقع لعرض رسالة الخطأ حينما تكون شاشة العرض صغيرة، فقد أظهرت النتائج أن اسفل حقل الادخال هو أفضل مكان. أما الشاشات الكبيرة فأفضل مكان لعرض رسالة الخطأ أسفل الدالة، اما فوق الدالة فقد أظهرت النتائج أنه أسوأ الأماكن لعرض رسالة الخطأ. أوضحت نتائج التجربة الأخيرة أن تظليل حقل الادخال عند حدوث الخطأ يساهم في تسهيل عملية تعبئة النماذج وتصحيح اخطاء الادخال.

ABSTRACT

Researchers have established a concrete and solid ground for web forms design and their usability testing since the nineties. However, since the dramatic shift towards devices that carries a varied set of properties, such as, smartphones and tablets, developers started to build forms to suit these devices. As a consequence, a radical shift in research and usability testing needed to strike as well. We attempt to work with Arabic forms that are designed particularly for mobile devices (smartphones to be exact). Our goal, was to come up with empirically tested usability guidelines for Arabic mobile forms through achieving a set of objectives. Firstly, we have investigated the effect of smartphones screen size on error rate. In other words, do smaller screens increase the number of errors made by the users during form filling tasks? Secondly, we have addressed the following question: what is the best location and timing for an error message to appear on smartphones forms? Finally, we have examined the use of graphics (highlighting the erroneous field) on form filling. The above issues as come to our knowledge never been addressed before for Arabic smartphones' forms. We have reached that conclusion after a detailed literature review. The above matters were measured in this dissertation in terms of efficiency, effectiveness, and satisfaction. Our experiments involved two main tasks: questionnaire distribution and form filling. Our findings indicated the existence of a negative relationship between screen size and number of errors made by the user in form filling, users are less likely to make errors when the screen is big. Our findings have shown that the best timing for an error message to appear is before submission (immediate feedback). This timing category involved less time needed to complete the forms and correct errors, less errors made by the users, and it was voted by the participants as the best timing. As for the best location for the small screen, our findings revealed that below text box is the best location for a message to appear, while for the big screen, below label is the best location for a message to be shown. Above label messages scored the worst which in turn implies that placing the error messages above the label is considered a bad news when it comes to error presentation. Our last experiment has shown that the use of graphics (highlighting the erroneous field) is very helpful. It captures the user attention instantly which implies the users ability to identify the location of the errors faster and much better.

Contents

List of Figures

List of Tables

Abbreviations

AL	**A**bove **L**able
ANOVA	**A**nalysis of **V**ariance
AS	**A**fter **S**ubmission
BL	**B**elow **L**able
BS	**B**efore **S**ubmission
BT	**B**elow **T**ext **B**ox
ECT	**E**rror **C**orrection **T**ime
EL	**E**rror **M**essage **L**ocation
EM	**E**rror **M**essage
ET	**E**rror Message **T**iming
HB	**H**ighlighted Text **B**ox
LB	**L**eft of **T**ext **B**ox
M	**M**ean
N	**N**umber of Participants
P	**P**robability
Per	**P**ercentage
PU	**P**op **U**p Message
RB	**R**ight of Text **B**ox
SD	**S**tandard **D**eviation

Chapter 1

Introduction

Researchers around the globe have given web forms usability a considerable amount of thought due to their prominence [13]. Moreover, the ground for web forms usability testing was set ages ago. A user form can be defined as a page that has boxes a user can type into [22]. The studies that have investigated and tested web forms usability are both myriad and miscellaneous. However, in the past few years mobile devices have supplanted desktops which involves in turn underwent research efforts for some facelifts [23]. Mobile devices are associated by a set of properties that might hamper users ability to complete their online tasks that include form filling [1]. In fact, the landscape associated with mobile devices keeps in spiking each and every day. This landscape involves the ability of accessing web on the go, due to their relative small screen real estate and other features [1]. Such set of features have made forms that are designed for desktops not suitable to use on mobile devices.

There is a friction of the previously conducted researches that has brought mobile forms to the spot light. However, these researches are very little and condensed. That perhaps might be attributed to the relative novelty of the use of mobile devices for online activities [46]. The research efforts varied from addressing design guidelines for forms to usability testing to identify wacky designs. The ISO group defined usability as: Extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use [47]. The studies will be discussed with more details latter in the literature review chapter.

In our work, we have concentrated our efforts on smartphones screen size and their implications on usability of Arabic forms. To be more specific we focused on their effect on the number of errors. Moreover, we have tested different locations and timing of error messages to identify the best amongst them. In addition, we have tested the effect of the use of graphics (highlighting erroneous field) within forms, to check if the highlighting erroneous field would lead for a better message affordance (message better and faster to be noticed). The previously mentioned issues as other researches have reported either never been tested or there are a scarcity and some limitations imposed on the studies that were directed towards addressing them. We hav used two samsung devices. The first one of a screen size 4.3 inches, this device is the one that we refer it as the small device in this document. The second one of a display size 5.5 inches, which we refer to it as the big device in this document.

1.1 Problem Definition

One of the most important issues to consider regarding forms is their usability since form filling is always perceived by the users as a tedious and boring task. A usable form encourages users to complete it without frustration and maybe some sort of satisfaction. Satisfied users in turn leads to less users turnover which will lead to a bigger audience for the given website. Now researchers have been studying and testing web forms for decades. There is an ample range of research in this sphere. However, when it comes to forms built for small devices, there is a seldom research that covers this area. Usability testing for mobile forms still considered brand new.

Form design and usability testing regarding small devices differ from those for web forms using desktops. That can be attributed to the wide range of characteristics of those devices. Those characteristics have made them portable yet they have imposed multiple restrictions on form design. These characteristics include small displays, multiple text- input methods and others. So, research efforts should concentrate on usability testing for mobile devices. Researchers have tested the usability of English web forms from various angles and as a result

they have come up with some empirically tested usability guidelines. However, for both Arabic and mobile forms the research is very little.

1.2 Objectives

The following is the list of the objectives we have seeked to achieve in this thesis:

- Investigating the effect of screen size on the number of errors made by the users in form filling process.
- Determining best location for the appearance of error messages for two smartphones with a different screen size.
- Determining best timing for the appearance of error messages for two mobiles with a different screen size.
- Investigating the effect of the use of graphics (highlight erroneous field) on the process of form filling. Does the use of graphics leads to less errors, better affordance (user notices the error message faster) better and faster form completion.

1.3 Research Questions

Research questions that we have attempted to address in this dissertation contains the following:

- What is the effect of smartphones screen size on form filling error rate?
- What is the best timing and location for the appearance of error messages in Arabic mobile form?
- How does the culture affect the location and timing of error messages?
- How does graphics (highlight erroneous field) affect user performance and form usability?

1.4 Contribution

- Identifying the effect of screen size on the number of errors.
- Identifying the best location to place an error message as well as best timing for its appearance.
- Identifying the possibility of increasing the efficiency of form filling through the use of graphics.
- Contributing to the usability of mobile Arabic forms, through facilitating form filling and identifying the best location and timing for displaying errors.

1.5 Thesis Organization

This section presents the outline of our thesis. Chapter 2 covers literature review. In this section, we have summarized some of the studies regarding web forms and testing their usability, impact of screen size related researches and error messages presentation. Methodology and experiments are discussed in Chapter 3. In this section we have presented the steps we have followed to accomplish our work, in addition to the experiments involved for doing so. Results are found in chapter 4. Discussion and conclusion of our work is given in Chapter 5. Finally, Chapter 6 covers the list of limitations and future work.

Chapter 2

Literature Review

A usable web form is attractive and enthusiasts users to complete it. Researchers and practitioners have been performing tests on web forms to ensure their usability. In this section, we will explain the definition of forms as it was stated by researchers, and the set of components that can be included in a form. Second, we will present some of studies that have looked in some depth to form usability testing. Finally, we will discuss studies related to small devices and effect of their screen size on activities performed, such as form filling.

2.1 Forms

Researchers have marked the componants of a web form. They include: input fields, labels, radio buttons, drop down lists, and others are called the components of a web form. Some other components that should be given some consideration are error messages. Issues related to error messages include, timing, location, wording and formatting. Previous researches, such as [50] [30] [13] [7], have illuminated the fact that error messages should be given a tremendous attention for reaching and achieving better form design and user experience. Some other researches have provided some sort of guidelines regarding error messages. For instance, error messages should be clear, short and comprehensive at the same time. They should convey the meaning with little used words as possible, and they should be easy to handle. Our work is related to error messages, and it is worth to mention that whether we are talking about web or mobile forms

they both have the same definition and components. However, regarding forms for mobiles we have to consider the issues and properties that accompanied with these small devices.

2.2 Forms usability testing

The matter usability of any product is always considered a priority, whether the product at hand is a device, an entire web site, or even a single web page that contains a web form. The term usability was defined in different ways by different groups of scholars. In this thesis we provide the ISO definition of usability. According to ISO, usability is: Extent to which a product can be used by specified users to achieve specified goals with effectiveness, efficiency, and satisfaction in a specified context of use [47]. It has been set a long time ago that efficiency, effectiveness, and satisfaction are three fundamental concepts for measuring and testing usability of any product including web forms.

As we have mentioned earlier web forms usability have been tested in respect to their design and their various components. Multiple form factors were tested, such as, label location, navigation, error messages timing and location, font size, font format, and others.

In [10] labels location was tested. Authors in [10] aim to combine two popular methods to evaluate web forms. These two methods are eye tracking and usability testing. Through the use of the above two methods, testers are able to evaluate forms in terms of their efficiency, accuracy, and satisfaction. The authors have concentrated their work on evaluating one of the form most important features which is label location. The main goal of this paper is to empirically evaluate some aspects of web forms as well as understanding and accentuate user cognitive process associated with the form layout. The latter is done through tracking the eye movement while interacting with the form. They have designed five versions of the same form, each differs than the other in terms of label location only as the following; Form 1: labels are left aligned. Form 2: labels are right aligned. Form 3: labels are above the input field. Form 4: labels are inside the input fields. Their results indicated that form 1 and 2 were easier to fill in comparison to the other forms and they were visually pleasing the most as well.

Users proclaimed that form 5 is not visually pleasing at all while form 4 was the toughest to fill.

In [46] researchers attempt to answer a crucial question: Does user culture (personal characteristics) affect the process of form filling or not? The previously conducted usability researches have neglected this question. So, the authors of [46] have decided to investigate this relationship. The results of their research and conducted experiments have proven that cultural dimensions affect the form filling process. Researchers suggested that the creation of a web site that can automatically adapt to suit users preferences and individual characteristics will improve their performance and reduce the amount of frustration while filling a web form.

Authors of [47] approached usability testing from a different angle. They were interested in shedding some light on how to represent the gathered data (in the log file) to the analysts and usability experts. In order to facilitate their mission in evaluating web sites usability. The authors have proposed a timeline-based system that is able to visualize the captured data. In a timeline system, every captured event or action is characterized by a name, timestamp, and other event related information. This system allows experts to accurately analyze user interactions with a web site and more easily analyze the collected data.

2.3 Screen size impact on user activities (from browsing to form filling)

The impact of screen size on usability has been investigated with mobile aspects (such as screen size) [34][37][42] and with mobile form factors (such as form componants) [18][19][24]. In [24] researchers have studied the impact of small-sized screens on information retrieval tasks. They wanted to address the following question: Do small screens affect web usability? The researchers have carried out an experiment to test that. The involved participants were asked to perform a set of tasks on a web based information system. Twenty volunteers have joined to participate in the test. Ten of them were asked to carry out the assigned tasks using a browser with its window display resolution set to 1074 X 768 pixels. The other ten were asked to perform the same task, but this time the browser display

capacity was set to 640 X 480 pixels. The results of their experiments have shown that those users who have used the smaller screen performed poorly. They were 50% less effective in their performance than those who used larger screens. The authors based on these results have provided a set of guidelines that should be considered during the design of a web site for small screens, for example, A direct research feature should be provided, and reduce scrolling as much as possible. This can be done by reducing the amount of content presented in the page and information are recommended to be organized using menus in fixed places.

This paper [34] has inspected the possible impact of screen size on learning that is based on videos. They seek to identify the issues that a video based learner might encounter when he or she is using a small device. The results of their research highlighted two interesting points. First, that users no matter the size of the screen used to view the video for learning have revealed a positive attitude towards the process (learning through video). Second, despite the users positive attitude there is some sort of a tradeoff when it comes to the effectiveness of the learning process. The smaller the device screen the less effective the process.

In [19], researchers have stated that desktop user interfaces (UI) do not suit small sized screens such as mobile screens. The previously conducted researches have designated multiple guidelines for redesigning the big (UI) so it fits the smaller devices, one of these guidelines regarding content (only a fraction of what is visible on the web site can be placed on mobile site). However, in some cases, the user is forced to navigate a large UI on a small screen. There are multiple ways that can be implemented to navigate in such case scenarios, this paper address some of these patterns: zooming, panning, and fisheye. This paper conducted a comparison study to identify the best amongst them. The results showed that fisheye was the best for performing navigational tasks while zooming was significantly better for monitoring tasks.

The researchers of [42] intended to explore the relationship between screen size of mobile devices and web usability. Their research orientate towards three main activities: navigation, searching, and form filling. The participants in the conducted experiments in this work were asked to interact with the same application, but on three different devices (Galaxy Ace, Galaxy Note, and Galaxy SII). The previous devices have different screen sizes. The results proved that larger screens lead to higher efficiency, usability, and effectiveness.

Mobile devices have small screens, retrieved information such as search results now have to be viewed through such devices. In [49] The researchers wanted to ensure if the screen-size of these devices affect the way search results are displayed, and if yes how to provide an appropriate summery for each result. Their investigation has approved that the screen size affect the summery of the viewed result. The presented summery on those devices has to be limited, specific, and representative as much as possible. They also suggested the use of contextual mechanisms to filter the retrieved search results. Contextual mechanisms involve considering the user current location, and environmental circumstances as well as user current information request and his long interest.

Some studies discussed the usability of mobile application [3][43][18]. Authors in [3] intended to test learning applications designed for smartphones. The applications were tested in terms of ease of use, user satisfaction, learnability, and attractiveness. The researchers distributed a questionnaire after requesting the participants to interact with mobile learning application designed by the authors based on a set of usability guidelines. The results indicated that users considered the proposed app user friendly and efficient to use.

In [43] the researchers aims to test the usability of a mobile website (library website), in an attempt to identify the issues that the user might encounter during the use of a library system that is conducted for mobile devices. The authors have combined two usability testing methodologies (field and laboratory experiments). Field tests enabled the researchers to test the library website on students native devices while laboratory tests enabled better control and recordings of test sessions. The experiments involved several kinds of devices (Android, iOS, Blackberry, and others). The recruited students were asked to interact with site through performing a set of tasks. Based on their results, the researchers came up with a set of recommendations to be considered when designing a library website for mobile phones, for instance, keep it simple and user friendly versions of error messages should be created.

Authors of [18] have investigated the usability of a flight booking app designed for mobile touch screens. Through their tests they have been able to designate the key usability-related problems in such an app. In addition, they have provided solutions for some of the identified problems. They have utilized a questionnaire and set up an interview for data capturing and gathering. An example of some

of the issues determined by the researchers and their proposed solution: Flight booking apps utilize wrong choices of icons: This kind of problems arise from the designers attempt to save space as much as they can. So, they use icons instead of data. That would help saving the screen space, but icons should carefully be chosen to convey the intended meaning to the users. There is another issue related to icons, that some companies try to uniform their icons, the ones used in their website and app. That cant be achieved due to mobile small screen size. The authors recommends the use of icons that are simple, clear, and easy to understand by the users.

2.4 Usability guidelines for mobile websites

Since the proliferation of mobile devices, a day after day users started to shift their online activities to be performed by those devices. Those devices can be used everywhere and anytime, and to add insult to injury, those devices are diverse and vary in size, shape, processing power, text input methods, and so many other features. Their surrogacy to desktops required changes in many activities including form design and their usability. Usability of mobile devices have been explored from different points of view.

Some form factors for mobile devices were tested and evaluated, such as, navigation [20]. The debate in [20] goes as the following: Design guidelines have recommended that long forms must be avoided, especially when it comes to small devices such as smartphones. However, in some situations designers find themselves in a loop, where they cant avoid long forms and at the same time they have to take into account that users view the process of filling up a form is tedious and cumbersome. In addition, todays mobile applications have the capability of providing complex functionality and to mention the limited real estate provided by mobiles. In such cases, researchers and practitioners have to think about effective solutions to sort out such dilemma, such as effective navigation. Authors constructed a usability testing study for comparing four different navigation methods or patterns (scrolling, tabs, menus, and collapsible fieldsets). The study was conducted to identify which one of these methods is more effective to use for navigating within a form that is viewed on a smartphone. They have

used the following measures to evaluate each pattern: memorability, usability, overview and subjective preferences. Their results have indicated that scrolling is a bad call to be used in navigation.

Some studies have recommended a set of usability guidelines for mobile design [3][47][14][32]. The constant increase in the production of smart phones has played as a great incentive for the creation of mobile applications. One of the main issues that contribute in the success of mobile apps is their usability. Usability guidelines for the web has been established since the 1990s. However, for mobile apps, researchers reported that no well-known and scientifically formed guidelines for mobile related apps. In this paper [47], authors aimed to create a catalogue that contains a set of well-defined usability guidelines for the sake of mobile apps development process. The authors reviewed the previously conducted studies in this field (mobile app usability guidelines). They have prepared a catalogue that contains a set of usability guidelines which are categorized as in table 1. The authors have conducted some experiments through the development of two mobile apps. They have used these cases to ensure the applicability of the guidelines reported in table 2.1. Based on the analysis of these results, the researchers reached to some conclusions:

1. Usability guidelines should be involved at an early stage of the apps' development process. By doing so, that will help to reduce the number of usability bugs, which implies shorter implementation time as well as enhanced overall result.

2. Most usability guidelines are applicable for both applications.

3. The usability guidelines catalogue should be kept updated and extended, in order to keep up with the rapid change in mobile and their apps development.

Limitations: The guidelines are applicable for both apps. However, more experiments should be established to ensure they are generally applicableand that they are sufficient and comprehensive.

Future work:

TABLE 2.1: Usability guidelines for mobile websites and applications [47]

Layout: (G1) Place content in the central part of the screen (G2) Avoid horizontal scrolling (G3) Arrange content vertically, avoid using tabs (G4) Locate login button on the top of the screen (G5) Make the search field visible and usable (G6) Use clear linear structure and avoid table structure of input fields (G7) Use no more than three buttons in a dialogue (G8) Position buttons in the middle or at the end of the dialogue (G9) Position buttons on the right side of the screen
Navigation: (G10) Make navigation menus as easy and as simple as possible (G11) Use one level navigation menus (G12) Avoid many options (G13) Make important pages reachable from the start page (G14) Provide a possibility to quickly navigate back to the start page (G15) Minimize the number of clicks needed to reach each page (G16) Use breadcrumbs to show current location (G17) Open external links in a new window, keep the current window unchanged (G18) Structure menus by topics and usage scenarios to reduce search (G19) Make titles, links and navigation items unambiguous
Design: (G20) Keep design simple (G21) Use a uniform design for different representations of the same product (e.g. website and mobile app) (G22) Emphasize important information (G23) Use easy identifiable and visible icons (G24) Use button size between 7mm and 10mm
Content: (G25) Use similar content and functionality for mobile version as for the desktop version (G26) Avoid long texts and use simple sentences (G27) Do not show much information on the start page (G28) Order content by its importance (G29) Provide automatic suggestions within the application (e.g. search autocomplete) (G30) Make user actions revertible to allow quick corrections (G31) Integrate confirmation dialogues for change and edit actions (G32) Prevent data loss by reminding users of unsaved changes (G33) Save the current application state when the application is minimized and restore again when the work is continued (G34) Automatically correct typing errors (G35) Optimize interface to correct imprecise touch control (G36) Implement haptic feedback on user actions (G37) Support common gestures to control apps (G38) Optimize entering passcode on mobile devices
Performance: (G39) Minimize loading times. Note: regular users are more tolerant to loading times than new users

1. More experiments to test the guidelines, involving other applications and other user groups.

2. The reported guidelines should be investigated by experts to be evaluated. They should be ranked by importance. In addition, least relevant guidelines should be identified and eliminated.

3. Guideline G9 should be investigated. Because it was applied in neither cases.

Most of the developed web sites were built to be used on large displays. However, after the proliferation of the small displays, the previously conducted web sites are not suitable for these devices. In addition, the design guidelines recommended for big screens stumbled, they can't be directly applied for the new smaller devices, such as, smartphones. Authors of [32] have intended to address this issue by providing a set of guidelines that will work for small screens. They also have provided a piece of code that can be injected in any web page, the result of executing that code is to report whether the user is using a desktop or mobile device to access the web, and if its a mobile device, then what type of device the user is using. Their set of guidelines include the following: keep it simple, simplify user input, scroll vertically and not horizontally, and build a multiple versions of website.

In [14] researchers discussed and tested the usability of commercial forms. They have orientated their research around form main attributes, such as, content, search, and navigation. Based on their research they have provided a set of guidelines that can be utilized for designing a commercial web site. Some of their proposed guidelines include: First, Avoid scrolling. If scrolling is needed then depend on vertical scrolling and avoid horizontal scrolling. Second, Use a flat hierarchy for structuring content and pages. Third, a Back button should be implemented in almost every page. The researchers recommended that developers should consider both expert and novice users. Those two categories of users tend to have different perception and behavior, as a result the design should suit both categories of users.

2.5 Error messages presentation

This section includes the studies that have been conducted for one of the most important form components, error messages. Researchers in [7] [37] [43] [45] evaluated error messages (location and/ or timing). In this work [45], researchers stated that there are several locations where error messages might be placed in a web form. Their goal is to investigate the impact of the location of error messages on form usability in terms of efficiency, effectiveness, and satisfaction. They have looked in depth to six different locations: right, left, above, below, top, and bottom. Their results indicated that the best location for an error message in a web form is to be placed to the right of the erroneous field. In [21] the usability of web forms was tested in terms of form completion time, error recognition time, the number of saccades, and error correction time. They have tested four different locations: to the right of input field, above it, top of the form, and the bottom of the form. They have used an eye tracking device to analyze users' interaction with the forms through eye movement. The eye tracking test has shown users were able to notice error messages faster when it was displayed on the right side of input field, which in turn implies less effort required by the user to understand the given message. In addition, users expressed that they prefer error messages that involve an example that will clarify to them how to fix the error.

Timing of error messages was investigated in [7] [45]. The authors of [45] compared three alternatives for showing error messages. These alternatives involved the timing of error messages (when the message will appear). They have investigated the appearance of error messages at the top of the page after completing the entire form, below the input field after completing the entire form, and below the input field after completing it. These options were studied in relation to effectiveness (which is defined by the authors as the rate by which users notice the error message as soon as it appears). The results showed that the third option was the worse among them. Users missed the message with a total of 40%. On the other hand, in the above two options error messages was missed less often by the users.

On the other hand, efforts in [7] are concentrated basically on the timing of errors. However, they have investigated several locations of error messages as it

is shown in figure 2.1. The figure shows two main options for an error message, whether to be displayed immediately or after the completion of the entire form. In addition, error message can be embedded inside form or to have the shape of a dialogue (pop up box). They have concluded that the best practice for showing an error message is to wait until the completion of entire form. They also recommended three possible ways to present the error message, the three options are all effective: Afterward, embedded in the form, all at once. Afterward, embedded, one by one. Afterward, in a dialogue, one by one.

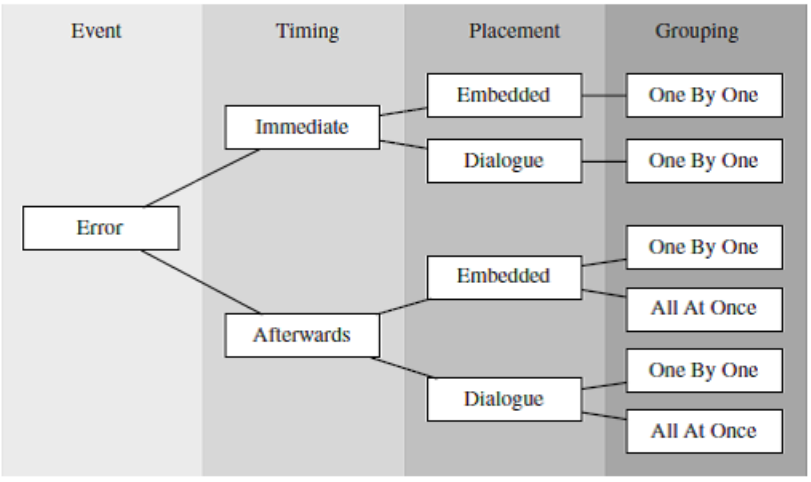

FIGURE 2.1: Main variables in error message presentation [7]

All of the previous studies have investigated error messages in respect to web forms. However, research regarding error messages for mobile forms have been given very little forethought. Moreover, researchers only investigated error messages for English forms. So, in this thesis, we carried our own research and experiments to investigate error messages regarding Arabic forms for smartphones. Table 2.2, 2.3, and 2.4 covers the limitations and future work of some of the previously conducted research.

Table 2.2: Summary of error messages related research

	Studied Attribute	Mobile or Web	Limitation	Future work
[1]	Error Messages	Web	Sample size is small. The study has neglected novice users.	Further work in inline feedback in revision mode. Exploring user behavior without eye tracking.
[7]	Error Messages	Web	Graphics are not involved.	Test to check if the users ignored error messages intentionally. Involve use of graphics. Study labels location.
[21]	Error Messages	Web	The involved users were all young adults and left to right readers. So, results may not be applicable to other groups of people. Novice users were neglected.	Future work should consider overcoming the limitations of this research.
[37]	Error Messages	Web	Specific locations were examined. Only soft errors involved. Statistics were imprecise and incomplete.	Test other locations. Test hard errors.
[46]	Error Messages	Web	Graphics are not involved. Short error messages only. Novice users were neglected. Users belong to the same culture.	Test longer forms. Test long error messages. Involve eye tracking. Involve graphics.

TABLE 2.3: Summary of screen size related work

	Studied Attrubute	Mobile or Web	Limitation	Future work
[18]	Screen size	Mobile	The experiments were conducted in a specific culture the findings may not be allowed to be generalized.	Further research to check if the findings are applicable to other cultures. Further research to identify the effect of screen resolution on video-based learning.
[19]	Screen size	Mobile	The provided guidelines only applicable to content designed specifically for small platforms display.	Investigating methods for automatic pages adaptation.
[43]	Screen Size	Mobile	Novice users were neglected.	Studying screen effect size on other tasks such as map navigation. Does effect of screen size change as the user become more familiar with the device. The bigger the size the better, to what point this positive relation holds.
[49]	Screen size	Mobile	The only involved experimental measure was accuracy.	Developing and employing the paper proposed system. Involving other experimental measures for the evaluation process.

Table 2.4: Summary of labels and navagation related work

	Studied Attribute	Mobile or Web	Limitation	Future work
[10]	Labels	Web	Sample size is small.	Involving other variables in form testing such as, form familiarity, complexity, label length, form length and others.
[13]	Navigation, menus and content structures	Web	Experimental site structure was simple and hierarchical, web sites may have semi-heretical or complex structure. The study used only common navigation methods and users are not allowed to use direct methods such as Home button.	Future studies will involve widen the experimental sites and the tasks involved. Other mobile devices must be involved as well.
[20]	Navagation	Mobile	Small sample size.	The study examined a scenario that involves social networks and form based user interface. The future work should involve other types of cases. Memorability of various navigational patterns should be examined further for long forms.

Chapter 3

Methodology

This chapter introduces the workflow and the steps followed to attain the goals and objectives that was drawn for this dissertation. Table 3.1 maps each research objective to related research questions, methods and instrument used to accomplish it. Before going any further, it is worth to mention that our methodology has started with a detailed literature review. In fact we have conducted two stages of literature review. We have begun the first stage the preliminary literature review in an attempt to identify the desired track of research that we want to be engaged in. The outcome of this stage is that the usability of Arabic smartphones' forms was elected as our primary research trajectory. The subsequent literature review was forged to identify the basic building blocks of our research. After an intensive search we have shifted error messages presentation and screen size as the core concepts for our research. Problem definition, research questions and objectives were articulated in posterior stages. They will not be disscussed any further, because they were elucidated in chapter one. The rest of the steps included in the methodology are discussed below. The following steps circle between the forms and tools used for testing and data gathering to the utlized methods and involved participants for the analysis of the collected data.

TABLE 3.1: Mapping objectives, reseach questions, methods and instruments

Objective	Research Question	Method	Instrument
Investigating the effect of screen size on the number of errors made by the users in form filling process.	What is the effect of smartphone screen size on form filling error rate?	Literature review and data collection through form filling, questionnaires and retrospective think aloud.	Form 4.1 To record the number of errors made by each user. Questionnaire One To collect demographic data.
Determining the best location and timing of error messages for two smartphones with a different screen size.	What is the best timing and location for the appearance of error messages in Arabic smartphones' forms?	Data collection through form filling, video recording, and questionnaires.	Form 4.2 till 4.10. To record error rate, time needed to fill the forms and correct the errors. Questionnaire Two To collect demographic and subjective ratings.
Investigating the effect of use of graphics (highlighting erroneous field) on form filling.	How does the graphics (highlighting erroneous field) affect user performance and forms usability?	Data collection through form filling, questionnaires and retrospective think aloud.	Form 4.11 To record error rates, form completion time, and error correction time. Questionnaire Two.

3.1 Form Design

We have designed 3 main forms, each suits the circumstances of each experiment. For example, for experiment 1, the form as it is shown in figure 3.1, contains a set of labels and input fields (text boxes only, no other types involved). The requested information are simple and easy to remember, the user doesnt need to take time to remember the answer, the answer already in his head. So, no time is required for thinking which might increase the time needed to complete the form. All of the involved fields are required with validation. We have placed some conditions (obstacles) on the requested data. The obstacles are summarized in table 3.2 and 3.3. The designed forms was tested by some individuals before the actual testing for this dissertation. We have asked ten random college students to test the designed forms for inspection purposes. We wanted them to check and report the faults they might find in the design. To test if it is comfortable to interact with the design and of course to check if our design live up to the real life forms' design. The students reported that our designs are good. They also reported that they are comfortable to interact with. In addition, they match to some level the real world forms.

Before going any further, we wanted to remind you that some of these obstacles doesn't exist in real life forms. However, that doesn't matter since that our main

FIGURE 3.1: Designed form for experiment 1.

goal to test the presentation of error messages as it was stated by [4]. The set of forms that we have designed and used in our experiments are shown in figures .

3.2 Testing and Data Collection Procedure

Researchers have pointed through their work that usability testing can be a quantitative testing, which deals with numbers and the analysis are done through some statistical and mathematical methods [24] [32]. On the other hand, some practitioners prefer to follow another route, which is the utilization of the qualitative approach in usability testing. This approach involves asking the participants about their experience with the tested object. It aims to collect some verbal description rather than values and numeric values. In other occasions, researchers tend to employ both approaches. In our dissertation both approches wrer harnessed. We have exploited questionnaires, retrospective think aloud (RTA) and video recording as the means to achieve that.

RTA refers to the process of asking the participants about their experience with

Table 3.2: Obstacles on the form fields for experiment 1

Field	Obstacles	Example
First Name	Arabic letters only.	
Last Name	English letters only. Uppercase.	SHARABATI
User Name	Starts with a letter. Followed by 4 characters. Ends with 9.	rtf_9
Date of Birth	D/M/Y Separated by /	8/9/1995
Address	Street number. Three numbers only.	123
Phone Number	Jawwal number only.	0596432567
Email	Hotmail or gmail.	rt%gy@gmail.com
Password	-	ty65r
Confirm Password	Matches password.	ty65r

Table 3.3: Obstacles on the form fields for experiments 2 and 3

Field	Obstacles	Example
First Name	English letters. Lowercase.	Ahmad
Date of Birth	D/M/Y Separated by /	8/9/1995
Phone Number	Jawwal number only.	0596432567
Email	Hotmail or gmail. In English.	rt%gy@gmail.com
Password	Starts with %. Followed by 4 letters. Ends with 2 digits.	%rtyu12
Confirm Password	Matches password.	%rtyu12

the tested object after completing the task at hand [2]. Some might argue that concurrent think aloud protocol (CTA) is more effective regarding data gathering, as the user can immediately express his thoughts, frustration, and doubts regarding the tasks they are handling and as a result they wont forget to report anything. However, in our case where the time to complete the task (complete form filling) is crucial, we can not depend on CTA, because it will increase the needed time to fill the form and the time needed to correct the errors. As a result we would receive inaccurate results. So, we have decided to pick RTA over CTA. We wanted the participants to voice their opinions, but without leading to affect the performance which in turn increases the time required to handle the task at hand.

FIGURE 3.2: Forms designed for experiment 1.

Each participant was asked to fill only one form. In other words, each participant belong to one and only one category. That is known to the researchers as the between-subjects approach [15]. This approch was picked over within-subject approach, to avoid the side effects that results from the latter. We didnt́ want the users to be affected by their experience that they have gained during filling the forms assigned to them. If the user was assigned more than one form he would commit less errors and take less time for next forms he fill and even if the order of the boxes were changed he would still manage to anticipate what he needs to do or how to avoid making errors and how to fix them quickly. This kind of experience would affect the results and would lead to generating less real data.

The data was collected using questionnaires and video recording. At the top of the each questionnaire, there is a statement that declares that any provided information by the user, written or recorded, is only used for scientific purposes,

FIGURE 3.3: Forms designed for experiment 2.

and any other types of usage the researcher takes full and complete responsibility for such behavior. The participants are asked for their consent by signing the questionnaire. Experiment one only involved the use of a questionnaire, a demographic questionnaire to be more specific and the number of errors made by the user while form filling was counted and written down directly after submission. Experiment two and three involved both recording and questionnaire distribution. In addition, two types of questionnaire were involved. The first one is a demographic questionnaire while the other one is related to error messages representation (location, timing, use of graphics and others) each one of them was rated by the students as strongly disagree, disagree, agree and strongly agree based on likert scale (1-4). The used questionnaires are provided in appendix of this dissertation.

FIGURE 3.4: EM appears above label

3.3 Appartatus

Android studio 2.3 was used to design the forms involved in the testing process. Java was used as the programming language to develop the forms while xml was used to design the forms. Figure 3.2, 3.3, 3.4, 3.5, 3.6, 3.7, 3.8, and 3.9 represent some of the resulted forms from the development process using android studio. Two Samsung devices were involved in the testing process, both run android as their operating system. The first one is Samsung Galaxy Star 2 plus of a screen size 4.3 inches (61.6% screen-to-body ratio), this device is the one that we refer it as the small device in this document. The second one is Samsung Galaxy A7 which we refer it as the big device in this document, of a display size 5.5 inches (74.3% screen-to-body ratio). The used operating systems, android KitKat 4.4.2 and marshmallow 6.0.1 respectively. For analyzing the collected data we have used SPSS 24.0 and alpha for all the tests was specified as 0.05.

FIGURE 3.5: EM appears below label

.png .png .png .png .png .png .png .pdf .pdf .pdf .pdf .pdf .pdf .pdf .jpg .jpg .jpg .jpg .jpg .jpg .jpg .jpg .mps .mps .mps .mps .mps .mps .mps .jpeg .jpeg .jpeg .jpeg .jpeg .jpeg .jpeg .jbig2 .jbig2 .jbig2 .jbig2 .jbig2 .jbig2 .jbig2 .jb2 .jb2 .jb2 .jb2 .jb2 .jb2 .jb2 .jb2 .PNG .PNG .PNG .PNG .PNG .PNG .PNG .PDF .PDF .PDF .PDF .PDF .PDF .PDF .JPG .JPG .JPG .JPG .JPG .JPG .JPG .JPEG .JPEG .JPEG .JPEG .JPEG .JPEG .JPEG .JBIG2 .JBIG2 .JBIG2 .JBIG2 .JBIG2 .JBIG2 .JBIG2 .JB2 .JB2 .JB2 .JB2 .JB2 .JB2 .JB2 .eps .eps .eps .eps .eps .eps .eps .eps .eps .eps .eps .eps .eps .eps .eps .eps .eps .eps .eps .eps

FIGURE 3.6: EM appears below text box

3.4 Measurments

The usability of mobile forms were tested using a number of metrics. Those metrics enfold efficiency, effectiveness, and satisfaction. As it was presented in [48] and [45], we can measure each one of them using various forms. We elucidate all of the used measures as the following:

1. Efficiency: we have measured efficiency by the time needed to complete the form (the moment the participant starts filling the form to the moment he or she submits the form successfully without errors). In addition, we have

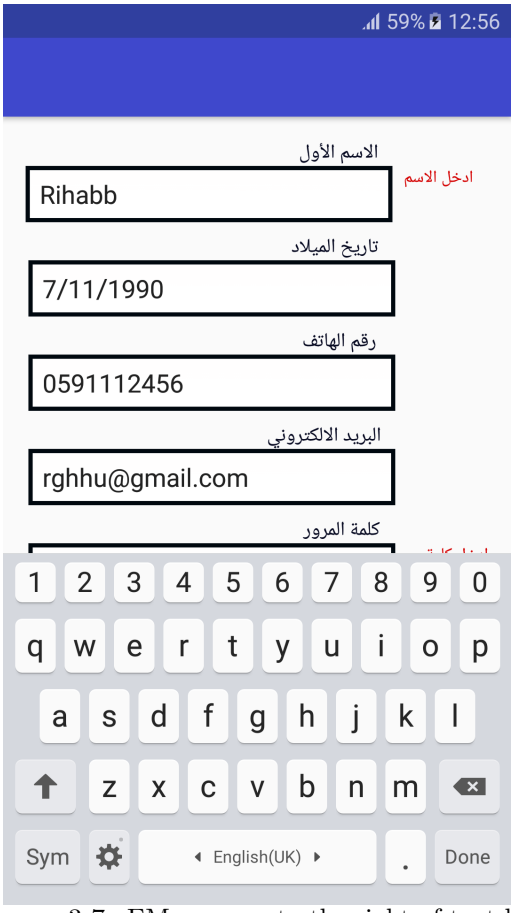

FIGURE 3.7: EM appears to the right of text box

measured efficiency by the time needed by the user to correct the errors that the subject made while filling the form at hand.

2. Effectiveness: In this dissertation, effectiveness is measured by error rate (number of first time errors, and number of consecutive errors made by the user).

3. Satisfaction: this one is measured by the participants ratings. The participants were asked to rate the presented error message regarding its location, timing or usage of graphics.

The above measures form the list of dependent variables in our thesis. The list of independent variable contains the method used to present the error message: its location, timing, or graphics (highlighting the erroneous input field).

FIGURE 3.8: EM appears to the left of text box

3.5 Experimental Design

We have conducted three main experiments. Each one is directed to address one of the main questions of this research. The sample of participants was selected from Palestine Polytechnic University. Only computer students (information technology, computer science, and computer system engineering) were involved. The reason behind our choices is that we wanted to ensure that all participants have dealt with computer and internet before, so their performance during experiments wont be affected by their experiences. If we have involved other majors or even literal participants, he or she may or may not have computer and internet experience before. Those will take much longer time in dealing with the task at hand, so we didnt want to take such a risk.

FIGURE 3.9: EM appears as a pop up message

3.6 Participants

We have conducted three experiments at Palestine Polytechnic University, the participants were randomly selected from the university. However, all participants are from the computer departments (Information technology, computers science, and computer systems engineering). All of them are undergraduate students. The number of recruited participants varied among the three experiments. Table 3.3 gives a summary of the total number of participants engaged in each experiment.

Before assigning any tasks to the students, they were asked to fill a questionnaire. The students were asked to answer questions related to gender, age, major, educational level and subject preferred device to browse the internet. In addition, the students were asked to rate their internet usage skills using smart phones, and how much time they spend on internet daily. The distributed questionnaire is provided in the appendix. We wanted the last three pieces of information to ensure that the students personal background does not affect their performance during the experiments.

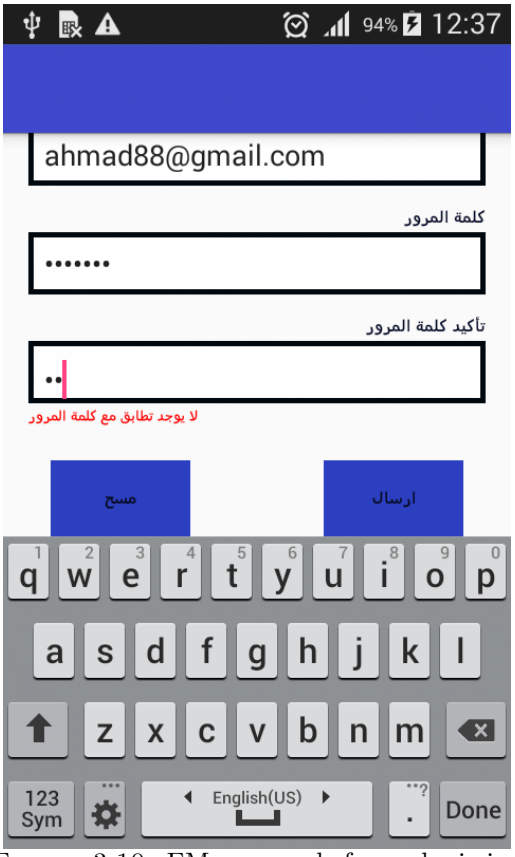

FIGURE 3.10: EM appears before submission

TABLE 3.4: Details of experiments environment.

	Experiment One	**Experiment Two**	**Experiment Three**
Studied Topic	Effect of Screen size on number of errors.	Best Timing and location of error messages.	Effects of graphics.
Used Device (2 Samsung Smartphones)	Galaxy Star 2 plus. Galaxy A7.	Galaxy Star 2 plus. Galaxy A7.	Galaxy A7.
Sample Size	100 student.	155 students.	66 students.
Time Required	10 days.	2 weeks.	5 days.

3.7 Experimental Procedure

Testing was carried out in Palestine Polytechnic University in April 2017. Each of the involved subjects was introduced to the device that will be used to fill the form. The experimental settings and purposes was explained to them as well before they start their assigned task. We explain the held experiments in more details below:

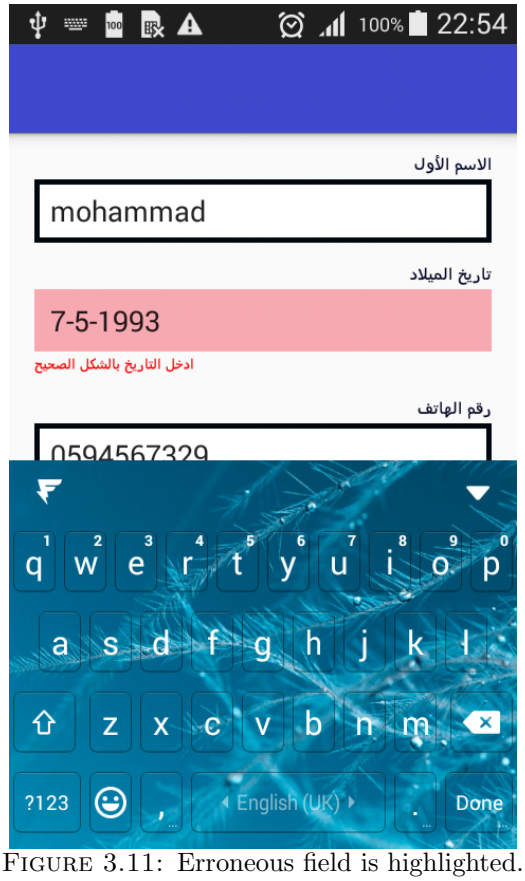

FIGURE 3.11: Erroneous field is highlighted.

- Experiment One: Effect of screen size on form filling

In experiment one where the effect of screen size on error rate is questioned, we have selected a hundred student. Each student is required to fill a questionnaire that contains some demographical aspects, such as age, gender, students educational level and others. The used questionnaire is available in the appendix. After that the participant was asked to fill the designed form. Two devices were involved, 50 students were assigned the form completion process using the small device, the other 50 were supposed to use the big device to fill the form with the required data. Each device were assigned an equal number of females and males, for instance, the big device was used by 25 females and 25 males.

Experiment one was targeting only the measurement of the number of errors that the user has made during form filling, in order to check if the number of errors differs among the two devices or not. The user fills the form and after the user finishes filling the assigned form, the user hits the submit button, and that is the end of it. Then the observer counts manually the number of errors made by the user.

- Experiment two: Identifying the best location and timing for displaying the errors

In experiment 2, new users are involved in this experiments. The user is required to correct the errors. If the user enters wrong data in a specific field, the user will be informed using an error message. The participant wont be allowed to leave until he submits the form successfully, and that happens after correcting all the errors. For the small device, we have tested two different forms of error messages. Embedded messages, the messages that appear near the erroneous field and separate error messages that appear as a pop up message. For the former, we have tested three different locations: above the label, below the label and below the input field, along with two different timing of error messages. For the previously mentioned locations plus pop up messages, 48 participants were involved. The participants were divided into groups, each group will test a form that requires the user to fill the same data but the error message is placed in a different location than the other forms assigned to other groups. Each group were assigned 12 participants (6 females, 6 males). For the big device, we have tested the previous locations, plus two other locations, messages placed to the left of the text box and to the right of the text box. After submitting the form successfully, the user was asked to fill a second questionnaire that is related to the location and timing of error messages.

For both devices we have tested two different timing of error messages. The first one involves the appearance of error messages after the user clicks the submit button, the other one involves the appearance of error messages immediately while the user is typing inside the text box. The process was videotaped and the observant latter analyzed each video for extracting the required information, such as, number of errors made by the user, type of the error, number of consecutive errors, time required to correct each error and the time required to complete the form successfully.

- Experiment three: Testing graphics (highlighting the erroneous fields)

Experiment 3 involved 66 new users. Each participant was asked to fill a questionnaire for demographic details. One device was involved, which is the big one. The goal of this experiment was to check the effect of the use of graphics on error rate and affordance (error better and faster noticed). The error field is highlighted with pink, since red is not favored, because we are considering

people with color blindness. 33 students were asked to test a form with only error messages (error message colored with red and was below text box and it appears after submission). The rest of them have tested the same form but with graphics involved. After form filling the participants were asked to fill a second questionnaire that is related to the presentation of error message. The process was videotaped and the observant latter analyzed each video for extracting the required information.

Chapter 4

Results and Analysis

This chapter presents the results of our experiments and their analysis. It includes the results and analysis of questionnaire and video recorded data.

4.1 Covariates

A hundred student participated in experiment one. Both genders were recruited, 50% females and 50% males. The average age was 21 years and SD was 1.84; the participants ages ranged between 18-27. As to the collage major a variety of disciplines were selected (IT, CS, and CSE). Educational level was among the following: a freshman (a student in his first year of college), a sophomore (a student in his second year of college), a junior (a student in his third year of college), a senior (a student in his fourth year of college), or a fifth year student (a student in his final year of college, if he is an engineer). The total number of participants included in each category are summarized below:

4.1.1 Experiment one: Effect of screen size on error rate

Small screen:

Freshman: 11 (22%). Sophomore: 10 (20%). Junior: 11 (22%). Senior: 11 (22%). Fifth year: 7 (14%).

Big screen:

Freshman: 9 (18%). Sophomore: 12 (24%). Junior: 10 (20%). Senior: 13 (26%). Fifth year: 6 (12%).

The students were asked to pick which device (mobile, PC) they prefer to use for surfing the internet:

Small screen PC: 14 (28%). Mobile: 36 (72%). Big screen PC: 21 (42%). Mobile: 29 (58%).

Individuals usage of internet:

Small screen:

Less than one hour: 4 (8%). 1 2 hours: 8 (16%). 3 5 hours: 26 (52%). 6 8 hours: 12 (24%).

Big Screen:

Less than one hour: 1 (2%). 1 2 hours: 12 (24%). 3 5 hours: 31 (62%). 6 8 hours: 6 (12%).

The analysis of the preferred device question have supported the importance of the directing the research towards mobile forms, since the number of users that uses mobile for internet browsing are more than those who uses PC.

A one way ANOVA test was performed to analyze the collected data. The analysis of our collected data regarding demographic characteristics, we have concluded that experimental samples dont vary. No significant differences found among the groups. As a result we have excluded the effects that might loom due to the demographic characteristics.

4.1.2 Experiment two: Identifying the best errors location and timing

In experiment two each screen (small and big screen) was tested separately. We have tested the location and timing of error messages. Once again at the beginning of the experiment we have distributed a demographic questionnaire. The small screen device was tested by a sample of 60 students. Tables 4.1, 4.2, 4.3, and 4.4 report the demographical data we have collected during small screen testing with respect to error message location, while Table 4.5 and 4.6 report the results of timing testing.

TABLE 4.1: Demographic Questionnaire, Small Screen, AL is Tested

Used Device: small screen			Tested Location: AL			
Attributes		N	Per	Range	M	SD
Gender	Female	6	50%			
	Male	6	50%			
Age				19 - 23	20	1.43
Major	IT	5	41.7%			
	CS	3	25%			
	CSE	4	33.3%			
Educational Level	Freshman	4	33.3%			
	Sophomore	4	33.3%			
	Junior	2	16.7%			
	Senior	1	8.3%			
	Fifth Year	1	8.3%			
Preferred Device	PC	5	41.7%			
	Mobile	7	58.3%			
Internet Skills	Bad	0	0	-	3.25	4
	Good	2	16.7%			
	Very good	5	41.7%			
	Excellent	5	41.7%			
Usage of Internet	Less than 1	2	16.7%			
	1 - 2 hours	1	8.3%			
	3 - 5 hours	6	50%			
	6 -8 hours	3	25%			

The big screen device was tested by a sample of 56 students. Tables 4.7, 4.8, 4.9, and 4.10 report the demographical data we have collected during big screen testing with respect to error message location, while Table 4.11 and 4.12 report the results of timing testing.

4.1.3 Experiment three: Effect of graphics on form filling

In experiment 3 we have tested two forms. The first one is BT which involves no graphics versus a second form that involves highlighting the erroneous input field. 66 participants were involved, 33 users for testing the forms with graphics and the rest for testing the form without graphics. The rest of the collected data is reported in Table 4.13 and 4.14.

TABLE 4.2: Demographic Questionnaire, Small Screen, BL is Tested

Used Device: small screen			Tested Location: BL			
Attributes		N	Per	Range	M	SD
Gender	Female	6	50%			
	Male	6	50%			
Age				18 - 25	19.58	2.02
Major	IT	3	25%			
	CS	4	33.3%			
	CSE	5	41.7%			
Educational Level	Freshman	6	50%			
	Sophomore	2	16.67%			
	Junior	3	25%			
	Senior	0	0			
	Fifth Year	1	8.3%			
Preferred Device	PC	3	25%			
	Mobile	9	75%			
Internet Skills	Bad	1	8.3%	-	3.17	.94
	Good	1	8.3%			
	Very good	5	41.7%			
	Excellent	5	41.7%			
Usage of Internet	Less than 1	0	0			
	1 - 2 hours	2	16.67%			
	3 - 5 hours	9	75%			
	6 -8 hours	1	8.3%			

4.2 Effect of screen size on error rate

4.2.1 Effectiveness

Effectiveness was measured by the number of errors made by the users in the process of form filling. As it has been determined by a one-way ANOVA that in experiment 1 there was a statistically significant difference between the two groups (small and big screen) regarding error rate. $F(1, 98) = 6.671$, $p = 0.011$. The analysis showed that users who filled the form using the small screen device have made much more errors than those who used the big screen device. Table 4.15 contains the mean and standard deviation of the experimental groups. The results proved that devices with small screens can lead to a higher error rate through form completion process. Most of the participants even expressed their thoughts regarding the matter, especially those who used the small device, they have voiced their opinions after finishing the task by saying: the device is small, its frustrating and derived me to make more errors during filling due its screen size. Figure 4.1 shows the results.

TABLE 4.3: Demographic Questionnaire, Small Screen, BT is Tested

Used Device: small screen			Tested Location: BT			
Attributes		N	Per	Range	M	SD
Gender	Female	6	50%			
	Male	6	50%			
Age				18 - 23	20.08	1.44
Major	IT	4	33.3%			
	CS	5	41.7%			
	CSE	3	25%			
Educational Level	Freshman	6	50%			
	Sophomore	7	58.3%			
	Junior	2	16.67%			
	Senior	0	0			
	Fifth Year	1	8.3%			
Preferred Device	PC	3	25%			
	Mobile	9	75%			
Internet Skills	Bad	0	0	-	3.08	.67
	Good	2	16.67			
	Very good	7	58.3%			
	Excellent	3	25%			
Usage of Internet	Less than 1	0	0			
	1 - 2 hours	2	16.67%			
	3 - 5 hours	9	75%			
	6 -8 hours	1	8.3%			

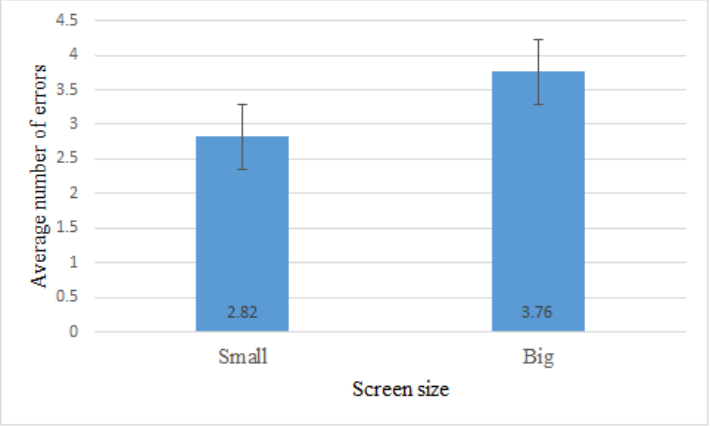

FIGURE 4.1: Error rate among the two tested devices

4.3 Identifying the best error location and timing

Experiment two was designed to test the location and timing of error messages in terms of efficiency, effectiveness, and satisfaction. The results and their analysis are summerized below in table 4.15, 4.16, 4.17, 4.18, 4.19, 4.20, 4.21, 4.22, 4.23, 4.24, 4.25, 4.26, 4.27, 4.28, 4.29, 4.30, 4.31, 4.32, 4.33, 4.34, 4.35, 4.36, 4.37, 4.38, and 4.39.

TABLE 4.4: Demographic Questionnaire, Small Screen, PU is Tested

Used Device: small screen			Tested Location: PU			
Attributes		N	Per	Range	M	SD
Gender	Female	6	50%			
	Male	6	50%			
Age				19 - 22	20.1	1.24
Major	IT	5	41.7%			
	CS	4	33.3%			
	CSE	3	25%			
Educational Level	Freshman	4	33.3%			
	Sophomore	2	16.67%			
	Junior	4	33.3%			
	Senior	2	16.67%			
	Fifth Year	0	0			
Preferred Device	PC	7	58.3%			
	Mobile	5	41.7%			
Internet Skills	Bad	0	0	-	2.75	.75
	Good	5	41.7%			
	Very good	5	41.7%			
	Excellent	2	16.67%			
Usage of Internet	Less than 1	1	8.3			
	1 - 2 hours	1	8.3%			
	3 - 5 hours	5	41.7%			
	6 -8 hours	5	41.7%			

4.3.1 Efficiency

In Table 4.16 we have reported the type of errors made regarding to each form (each form differs in the way the error message was located) that have been tested using the small device. The first column in the table represents the type of error, for example, first name in table 16 represents the errors made in first name field. The second column contains the mean and standard deviation of the time needed to correct each error. In other words, all of the users who have tested the form where the error message was positioned above the label (AL), plus those who have made errors while entering the first name.

A one-way ANOVA test was conducted to determine if there are significant differences between the tested groups (AL, BL, BT, and PU). Each error was tested separately from the rest of the errors. Time taken to correct each error was the dependent variable. One-way ANOVA has confirmed that no significant difference between the groups, $F(3, 44) = 2.587$, $p = 0.065$. So, time taken to correct the errors made in first name field is close across the different tested categories. Time taken to correct date of birth errors is statistically not significant among

TABLE 4.5: Demographic Questionnaire, Small Screen, BS is Tested

Used Device: small screen			Tested Timing: BS			
Attributes		N	Per	Range	M	SD
Gender	Female	6	50%			
	Male	6	50%			
Age				19 - 23	20.8	1.13
Major	IT	3	25%			
	CS	4	33.3%			
	CSE	5	41.7%			
Educational Level	Freshman	2	16.67%			
	Sophomore	4	33.3%			
	Junior	4	33.3%			
	Senior	0	0			
	Fifth Year	2	16.67%			
Preferred Device	PC	9	75%			
	Mobile	3	25%			
Internet Skills	Bad	0	0	-	2.76	.78
	Good	6	50%			
	Very good	4	33.3%			
	Excellent	2	16.67%			
Usage of Internet	Less than 1	0	0			
	1 - 2 hours	2	16.67%			
	3 - 5 hours	1	8.3%			
	6 -8 hours	9	75%			

the groups, F(3, 44) = 1.879, p = 0.147. The same goes for the time taken to correct phone, email, password and password confirmation as well. Results are presented in figure 4.2.

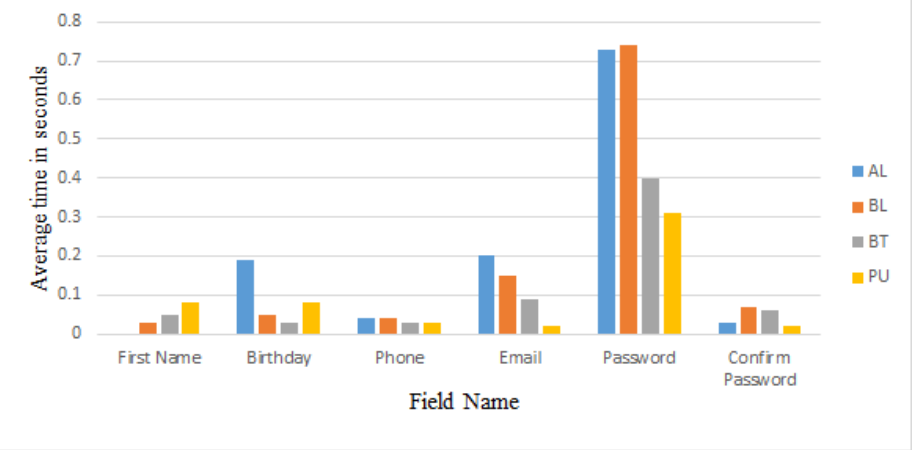

FIGURE 4.2: Average time to correct each type of errors using small device.

In table 4.17 the mean and standard deviation was calculated for all errors made by the participants who have tested the same form. Participants who were involved in testing the location of error messages (AL, BL, BT, and PU) using

TABLE 4.6: Demographic Questionnaire, Small Screen, AS is Tested

Used Device: small screen			Tested Timing: AS			
Attributes		N	Per	Range	M	SD
Gender	Female	6	50%			
	Male	6	50%			
Age				18 - 23	20.08	1.44
Major	IT	4	33.3%			
	CS	5	41.7%			
	CSE	3	25%			
Educational Level	Freshman	6	50%			
	Sophomore	7	58.3%			
	Junior	2	16.67%			
	Senior	0	0			
	Fifth Year	1	8.3%			
Preferred Device	PC	3	25%			
	Mobile	9	75%			
Internet Skills	Bad	0	0	-	3.08	.67
	Good	2	16.67			
	Very good	7	58.3%			
	Excellent	3	25%			
Usage of Internet	Less than 1	0	0			
	1 - 2 hours	2	16.67%			
	3 - 5 hours	9	75%			
	6 -8 hours	1	8.3%			

the small screen have corrected the errors in average of: 15.61, 14.99, 4.01, and 6.96 seconds respectively. The longest time to correct the errors was associated with error messages presented above the label. The shortest time was associated with error messages displayed below the text box. Differences between group means were diffidently significant as it was shown by one-way ANOVA test, $F(3, 44) = 6.080$, $p = 0.001$. As a result of this significance between the groups, we have continued the analysis using LSD to determine the group that gave the best result. LSD post hoc analysis revealed that some of the groups vary significantly than others, in other words there is a statistically significant differences between some of the groups as the following: AL and BT where $p = 0.001$, AL and PU where $p = 0.012$, BL and BT where $p = 0.002$, BL and PU where $p = 0.019$. However, we have observed that LSD post-hoc analysis revealed no significance what so ever when it comes to AL and BL where $p = 0.852$, as well as BT and PU where $p = 0.392$. The results are shown in figure 4.3.

In Table 4.18 we have reported mean and standard deviation for error correction time for the big device. A one-way ANOVA test was conducted to determine

TABLE 4.7: Demographic Questionnaire, Big Screen, AL is Tested

Used Device: big screen			Tested Location: AL			
Attributes		N	Per	Range	M	SD
Gender	Female	4	50%			
	Male	4	50%			
Age				20 - 23	21.13	0.99
Major	IT	3	37.5%			
	CS	2	37.5%			
	CSE	3	25%			
Educational Level	Freshman	0	0			
	Sophomore	0	0			
	Junior	2	25%			
	Senior	5	62.5%			
	Fifth Year	1	12.5%			
Preferred Device	PC	3	37.5%			
	Mobile	5	62.5%			
Internet Skills	Bad	0	0	-	3	.76
	Good	2	25%			
	Very good	4	50%			
	Excellent	2	25%			
Usage of Internet	Less than 1	0	0			
	1 - 2 hours	1	12.5%			
	3 - 5 hours	3	37.5%			
	6 -8 hours	4	50%			

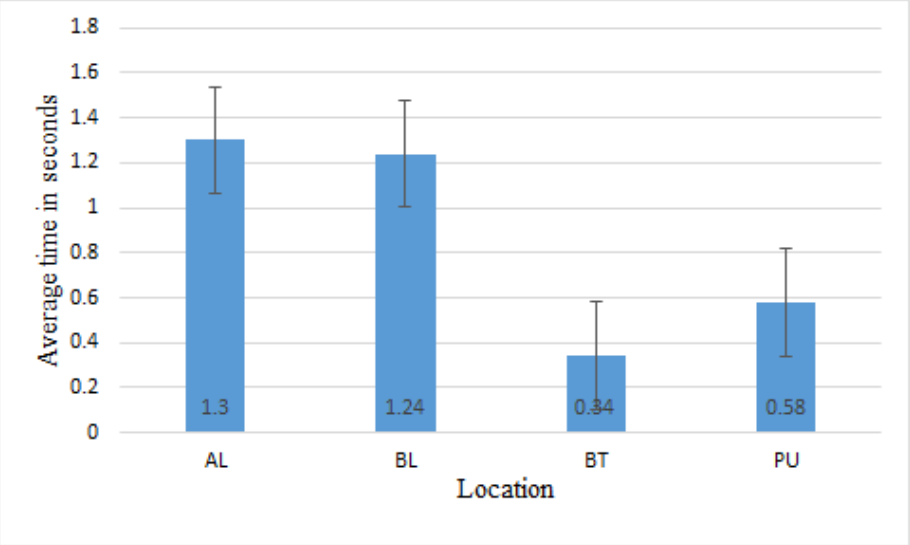

FIGURE 4.3: Time to correct errors of location tested groups using small device.

if there are significant differences between the tested groups (AL, BL, BT, PU, LB, and RB). Each error was tested separately from the rest of the errors. Time taken to correct each error was the dependent variable. One-way ANOVA has confirmed that no significant differences between the groups for all fields. The time associated with the tested groups are given by figure 4.4.

TABLE 4.8: Demographic Questionnaire, Big Screen, BL is Tested

Used Device: big screen			Tested Location: BL			
Attributes		N	Per	Range	M	SD
Gender	Female	4	50%			
	Male	4	50%			
Age				18 - 23	19.88	1.73
Major	IT	3	37.5%			
	CS	3	37.5%			
	CSE	2	25%			
Educational Level	Freshman	3	37.5%			
	Sophomore	2	25%			
	Junior	2	25%			
	Senior	0	0			
	Fifth Year	1	12.5%			
Preferred Device	PC	7	87.5%			
	Mobile	1	12.5%			
Internet Skills	Bad	0	0	-	3.13	.83
	Good	2	25%			
	Very good	3	37.5%			
	Excellent	3	37.5%			
Usage of Internet	Less than 1	1	12.5%			
	1 - 2 hours	2	25%			
	3 - 5 hours	4	50%			
	6 -8 hours	1	12.5%			

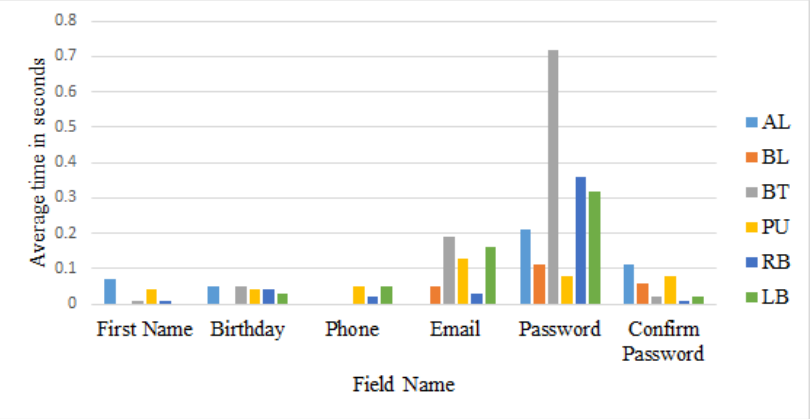

FIGURE 4.4: Average time to correct each type of errors using big device.

In table 4.19 the mean and standard deviation was calculated for all errors made by the participants who have tested the same form. Participants who were involved in testing the location of error messages (AL, BL, BT, LT, PU and RT) using the big screen have corrected the errors in average of: 4.27, 3.25, 7.91, 4.39, 2.14 and 3.65 seconds respectively. The longest time to correct the errors was associated with error messages presented below the text box. The shortest time was associated with error messages displayed as a pop up message.

TABLE 4.9: Demographic Questionnaire, Big Screen, BT is Tested

Used Device: big screen			Tested Location: BT			
Attributes		N	Per	Range	M	SD
Gender	Female	4	50%			
	Male	4	50%			
Age				19 - 24	21	1.69
Major	IT	3	37.5%			
	CS	2	25%			
	CSE	3	37.5%			
Educational Level	Freshman	2	25%			
	Sophomore	1	12.5%			
	Junior	2	25%			
	Senior	2	25%			
	Fifth Year	1	12.5%			
Preferred Device	PC	2	25%			
	Mobile	6	75%			
Internet Skills	Bad	0	0	-	3.13	.83
	Good	2	25%			
	Very good	3	37.5%			
	Excellent	3	37.5%			
Usage of Internet	Less than 1	1	12.5%			
	1 - 2 hours	2	25%			
	3 - 5 hours	4	50%			
	6 -8 hours	1	12.5%			

Differences between group means were definitely significant as it was shown by one-way ANOVA test, $F(5, 42) = 2.794$, $p = 0.029$. As a result of this significance between the groups, we have continued the analysis using LSD to determine the group that gave the best result. LSD post hoc analysis revealed that some of the groups vary significantly than others, in other words there is a statistically significant differences between some of the groups as the following: AL and BT where $p = 0.036$, BL and BT where $p = 0.002$, BT and PU where $p = 0.003$, BT and RT where $p = 0.027$. However, we have observed that LSD post-hoc analysis revealed no significance whatsoever when it comes to AL and BL, AL and LT, AL and PU, AL and RT, BL and LT, BL and PU, BL and RT, BL and LT, LT and PU as well as LT and RT. Figure 4.5 summerizes the collected data.

In Table 4.20 we have reported the type of errors made regarding to each form (timing of error message was different) that have been tested using the small device. A one-way ANOVA test was conducted to determine if there are significant differences between the tested groups (AS, BS). Each input field was tested separately from the rest of the errors. Time taken to correct each error

TABLE 4.10: Demographic Questionnaire, Big Screen, PU is Tested

Used Device: big device			Tested Location: PU			
Attributes		N	Per	Range	M	SD
Gender	Female	4	50%			
	Male	4	50%			
Age				19 - 22	20.1	1.24
Major	IT	1	12.5%			
	CS	4	50%			
	CSE	3	37.5%			
Educational Level	Freshman	3	37.5%			
	Sophomore	2	25%			
	Junior	2	25%			
	Senior	1	12.5%			
	Fifth Year	0	0			
Preferred Device	PC	2	25%			
	Mobile	6	75%			
Internet Skills	Bad	1	12.5%	-	2.12	.83
	Good	1	12.5%			
	Very good	3	37.5%			
	Excellent	3	37.5%			
Usage of Internet	Less than 1	2	25%			
	1 - 2 hours	3	37.5%			
	3 - 5 hours	2	25%			
	6 -8 hours	1	12.5%			

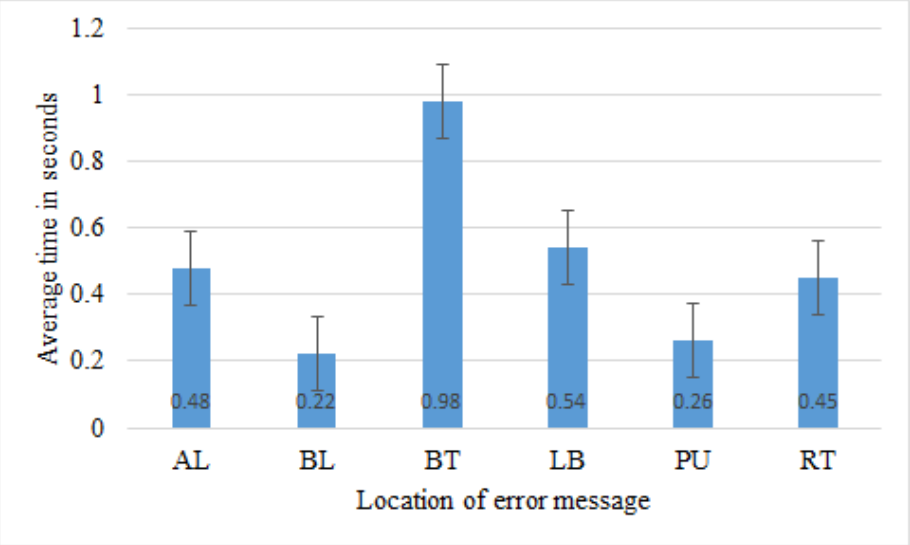

FIGURE 4.5: Average time to correct all errors using big device.

was the dependent variable. One-way ANOVA has confirmed that no significant differences between the groups (AS, BS) for all the fields. The time needed to correct the errors for each tested category is given by figure 4.6.

Table 4.21 shows the mean and standard deviation for experimented groups (AS, BS) using the small device. Users who submitted the form after correcting the

TABLE 4.11: Demographic Questionnaire, Big Screen, BS is Tested

Used Device: big device			Tested Timing: BS			
Attributes		N	Per	Range	M	SD
Gender	Female	16	51.5%			
	Male	17	48.5%			
Age				19 - 24	20.4	1.19
Major	IT	13	39.4%			
	CS	7	21.2%			
	CSE	10	30.3%			
Educational Level	Freshman	7	21.2%			
	Sophomore	9	37.5%			
	Junior	11	27.3%			
	Senior	3	9.1%			
	Fifth Year	3	9.1%			
Preferred Device	PC	11	27.3%			
	Mobile	22	66.7%			
Internet Skills	Bad	3	9.1%	-	2.54	.78
	Good	14	42.2%			
	Very good	12	36.7%			
	Excellent	4	12.1%			
Usage of Internet	Less than 1	2	6.1%			
	1 - 2 hours	3	9.1%			
	3 - 5 hours	12	36.4%			
	6 -8 hours	16	48.5%			

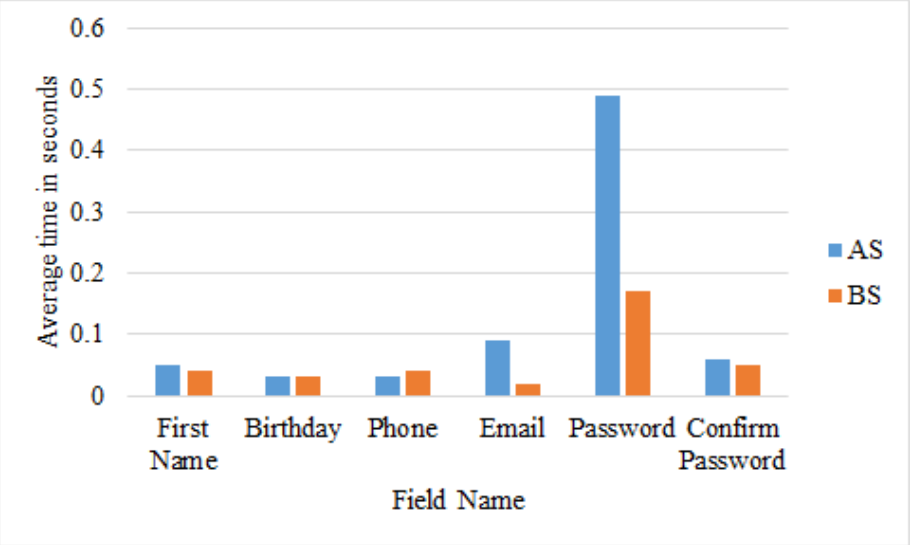

FIGURE 4.6: Average time to correct each type of errors using small device, timing groups.

errors have an average time of 5.1 seconds while the other group has corrected the errors with an average time of 9. 27 seconds. That result indicates that users who corrected the errors before submission took less time correcting the errors made while filling the required form. However, a one-way ANOVA suggested that the two groups dont vary greatly, $F(1,22) = 2.219$, $p = 0.151$. In other

TABLE 4.12: Demographic Questionnaire, Big Screen, AS is Tested

Used Device: big screen			Tested Timing: AS			
Attributes		N	Per	Range	M	SD
Gender	Female	16	48.5%			
	Male	17	51.5%			
Age				18 - 24	20.24	1.62
Major	IT	12	36.36%			
	CS	12	36.36%			
	CSE	9	27.27%			
Educational Level	Freshman	8	24.24%			
	Sophomore	11	33.33%			
	Junior	8	24.24%			
	Senior	4	12.12%			
	Fifth Year	2	6.06%			
Preferred Device	PC	12	36.36%			
	Mobile	21	63.64%			
Internet Skills	Bad	9	27.27%	-	2.72	1.06
	Good	6	18.18%			
	Very good	12	36.36%			
	Excellent	9	27.27%			
Usage of Internet	Less than 1	5	15.15%			
	1 - 2 hours	7	21.21%			
	3 - 5 hours	16	48.48%			
	6 -8 hours	5	15.15%			

words, no statistically significant differences between the two groups regarding total time to correct the errors made by the participants. Figure 4.7 summerizes the results.

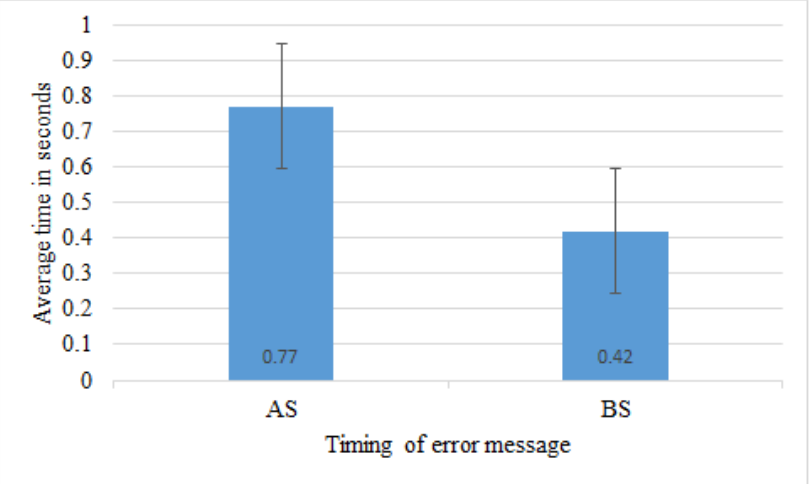

FIGURE 4.7: Time to correct all errors of timing tested groups using small device.

In Table 4.22 we have reported the type of errors made regarding to each form (timing of error message was different) that have been tested using the big device.

TABLE 4.13: Demographic Questionnaire, HB is Tested

Used Device: big screen			Tested Method: HB			
Attributes		N	Per	Range	M	SD
Gender	Female	17	51.5%			
	Male	16	48.5%			
Age				18 - 24	19.85	1.56
Major	IT	9	27.2%			
	CS	14	42.4%			
	CSE	10	30.3%			
Educational Level	Freshman	11	33.3%			
	Sophomore	8	24.2%			
	Junior	6	18.2%			
	Senior	7	21.2%			
	Fifth Year	1	3.03%			
Preferred Device	PC	21	63.6%			
	Mobile	12	36.4%			
Internet Skills	Bad	6	18.2%	-	2.67	1.02
	Good	6	18.2%			
	Very good	14	42.4%			
	Excellent	7	21.2%			
Usage of Internet	Less than 1	6	18.18%			
	1 - 2 hours	3	9.09%			
	3 - 5 hours	12	36.36%			
	6 -8 hours	12	36.36%			

A one-way ANOVA test was conducted to determine if there are significant differences between the tested groups (AS, BS). Each input field was tested separately from the rest of the errors. Time taken to correct each error was the dependent variable. One-way ANOVA has confirmed that no significant differences between the groups (AS, BS) for all the fields. Figure 4.8 presents the gathered data for timing tested groups.

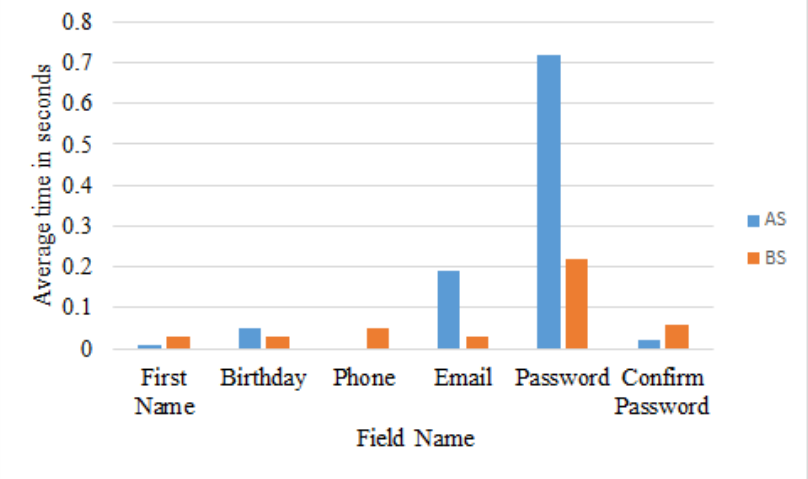

FIGURE 4.8: Time to correct errors of timing tested groups using big device.

Table 4.14: Results of Demographic Questionnaire, Without graphics

Used Device: Big screen			Tested Method: Without graphics			
Attributes		N	Per	Range	M	SD
Gender	Female	16	48.5%			
	Male	17	51.5%			
Age				18 - 24	20.24	1.62
Major	IT	12	36.36%			
	CS	12	36.36%			
	CSE	9	27.27%			
Educational Level	Freshman	8	24.24%			
	Sophomore	11	33.33%			
	Junior	8	24.24%			
	Senior	4	12.12%			
	Fifth Year	2	6.06%			
Preferred Device	PC	12	36.36%			
	Mobile	21	63.64%			
Internet Skills	Bad	9	27.27%	-	2.72	1.06
	Good	6	18.18%			
	Very good	12	36.36%			
	Excellent	9	27.27%			
Usage of Internet	Less than 1	5	15.15%			
	1 - 2 hours	7	21.21%			
	3 - 5 hours	16	48.48%			
	6 -8 hours	5	15.15%			

*Same as AS results.

Table 4.15: Error Rate in experiment 1

Screen size	N	M	SD
Big	50	2.82	1.870
Small	50	3.76	1.768
Total	100	3.29	1.871

Table 4.16: Time to correct the errors, small screen, location is tested

Error	AL		BL		BT		PU	
Message	M	SD	M	SD	M	SD	M	SD
First Name	.00	.00	.03	.08	.05	.08	.08	.09
Birthday	.19	.31	.05	.10	.03	.07	.08	.13
Phone Number	.04	.11	.04	.098	.03	.095	.03	.06
Email	.20	.43	.15	.21	.09	.16	.02	.06
Password	.73	.81	.74	.71	.40	.56	.31	.56
Confirm Password	.03	.08	.07	.13	.06	.08	.02	.06

*Small screen. Error message Location is tested.

Table 4.23 shows the mean and standard deviation for experimented groups (AS, BS) using the big device. The tested independent variable is the timing of error messages. Users who submitted the form after correcting the errors have an average time of 10.89 seconds while the other group has corrected the errors with an average time of 21.9 seconds. That result indicates that users

Table 4.17: Time taken to correct all errors

Method	N	M	SD
AL	12	1.3008	.84583
BL	12	1.2492	.80747
BT	12	.3417	.37868
PU	12	.5800	.55829
Total	48	.8679	.77678

*Small screen. Error message Location is tested.

Table 4.18: Time to correct the errors, big screen, location is tested

Error Message	AL (SD)	BL (SD)	BT (SD)	PU (SD)	RB (SD)	LB (SD)
First Name	.07 .17	.00 (.00)	.01 (.04)	.04 (.01)	.01 (.04)	.00 (.00)
Birthday	.05 (.10)	.00 (.00)	.05 (.07)	.04 (.08)	.04 (.07)	.03 (.08)
Phone Number	.00 (.00)	.00 (.00)	.00 (.00)	.05 (.02)	.02 (.08)	.05 (.13)
Email	.00 (.00)	.05 (.0995)	.19 (.398)	.13 (.35)	.03 (.07)	.16 (.06)
Password	.21 (.17)	.11 (.15)	.72 (.77)	.08 (.15)	.36 (.51)	.32 (.41)
Confirm Password	.11 (.16)	.06 (.095)	.02 (.07)	.08 (.14)	.01 (.02)	.02 (.01)

*Big screen. Error message Location is tested.

Table 4.19: Time taken to correct all errors in experiment 2, big screen

Method	N	M	SD
AL	8	.4875	.48186
BL	8	.2213	.07318
BT	8	.9887	.76005
LT	8	.5488	.36451
PU	8	.2675	.34028
RT	8	.4563	.47377
Total	48	.4950	.50567

*Big screen. Error message Location is tested.

who corrected the errors before submission took less time correcting the errors made while filling the required form. Moreover, a one-way ANOVA suggested that the two groups vary, $F(1, 64) = 6.85$, $p = 0.011$. In other words, there is a statistically significant difference between the two groups regarding total time to correct the errors made by the participants. Figure 4.9 shows the average time for the timing tested groups using the big device. The data are provided by figure 4.9. Table 4.24 carries data related to the total time needed to complete the entire form. The mean and standard deviation is listed for the tested groups (AL, BL, BT, and PU). The participants have used the small device to fill the

TABLE 4.20: Time to correct the errors, Small screen, timing is tested

Error	AS		BS	
Message	M	SD	M	SD
First Name	.05	.08	.04	.07
Birthday	.03	.07	.03	.08
Phone Number	.03	.095	.04	.08
Email	.09	.16	.02	.07
Password	.49	.57	.17	.42
Confirm Password	.06	.08	.05	.08

TABLE 4.21: Time to correct all errors, timing is tested

Method	N	M	SD
AS	12	.7725	.64944
BS	12	.4250	.48105
Total	24	.5988	.58642

*Small screen. Error message Timing is tested.

TABLE 4.22: ECT in experiment 2, Timing using big device

Error	AS		BS	
Message	M	SD	M	SD
First Name	.01	.04	.03	.05
Birthday	.05	.07	.03	.06
Phone Number	.00	.00	.05	.06
Email	.19	.398	.03	.07
Password	.72	.77	.22	.45
Confirm Password	.02	.07	.06	.10

TABLE 4.23: Time to correct all errors, big screen, timing tested

Method	N	M	SD
AS	33	.6636	.6196
BS	33	.3300	.3903
Total	66	.4968	.5406

*Big screen. Error message Timing is tested.

TABLE 4.24: Time complete the form, Small screen, Location is tested

Method	N	M	SD
AL	12	3.0492	1.1084
BL	12	2.4900	.87483
BT	12	2.8133	.82876
PU	12	2.3333	1.26664
Total	48	2.6715	1.04022

*Small screen. Error message Location is tested.

forms. The users who filled the form with the error messages displayed above the label (AL) completed the form with a total time 36.59 seconds, which is the longest time among the groups. The shortest time was delivered by those who have completed the form (PU) which is 28 seconds. However, one-way ANOVA test have revealed no significant differences between the tested groups regarding

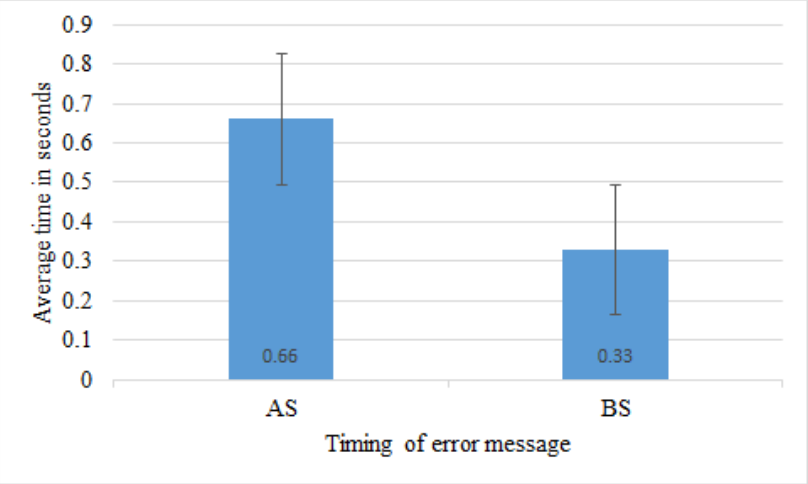

FIGURE 4.9: Time to correct each type of errors of timing tested groups using big device.

form completion time, F(3, 44) = 1.158, p = 0.337. Figure 4.10 show the time related data for small device location tested groups. Table 4.25 carries data

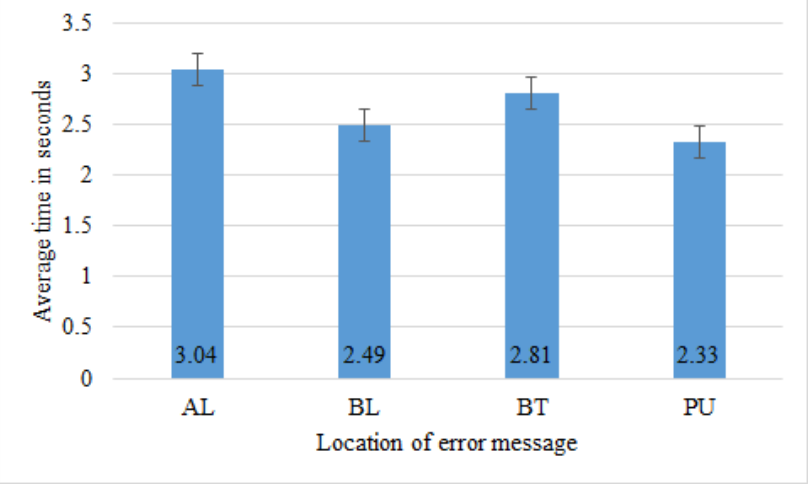

FIGURE 4.10: Average completion time of location tested groups using small device.

TABLE 4.25: Time to complete the form, big screen, Location is tested

Method	N	M	SD
AL	8	1.9575	.54374
BL	8	1.3175	.36745
BT	8	1.9313	.65856
LB	8	1.8537	.68623
PU	8	2.1538	.50068
RB	8	1.6912	.41995
Total	48	1.8175	.57684

*Big screen. Error message Location is tested.

related to the total time needed to complete the entire form. The mean and standard deviation is listed for the tested groups (AL, BL, BT, LB, PU, and

RB). The participants have used the big device to fill the forms. The users who filled the form with the error messages displayed to the left of text box (LB) completed the form with a total time 17.23 seconds, which is the longest time among the groups. The shortest time was delivered by those who have completed the form (BL) where the error message was displayed below the label which is 14 seconds. However, one-way ANOVA test have revealed no significant differences between the tested groups regarding form completion time, $F(5, 42) = 2.249$, $p = 0.067$. Data are presented by figure 4.11.

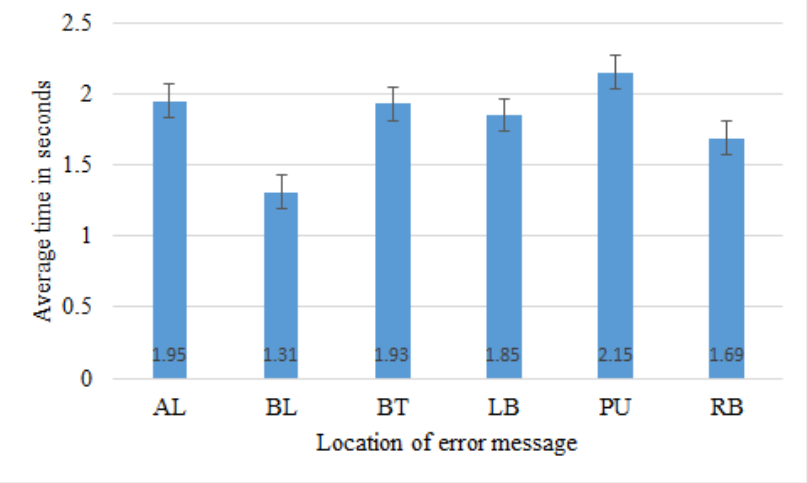

FIGURE 4.11: Average completion time of location tested groups using big device.

TABLE 4.26: Time complete the form, small screen, timing is tested

Method	N	M	SD
AS	12	2.8133	.82876
BS	12	1.8117	.77248
Total	24	2.3125	.93575

*Small screen. Error message Timing is tested.

Table 4.26 carries data related to the total time needed to complete the entire form. The mean and standard deviation is listed for the tested groups (AS, BS). The participants have used the small device to fill the forms. The users of AS forms have a total time of 33.76 seconds, which is longer than BS group which is 21.74 seconds. A one-way ANOVA test have revealed that there is a statically significant difference between the tested groups regarding form completion time, $F(1, 22) = 9.380$, $p = 0.006$. That implies that displaying error messages before submitting leads to a faster form completion time. AS and BS form completion time provided by figure 4.12.

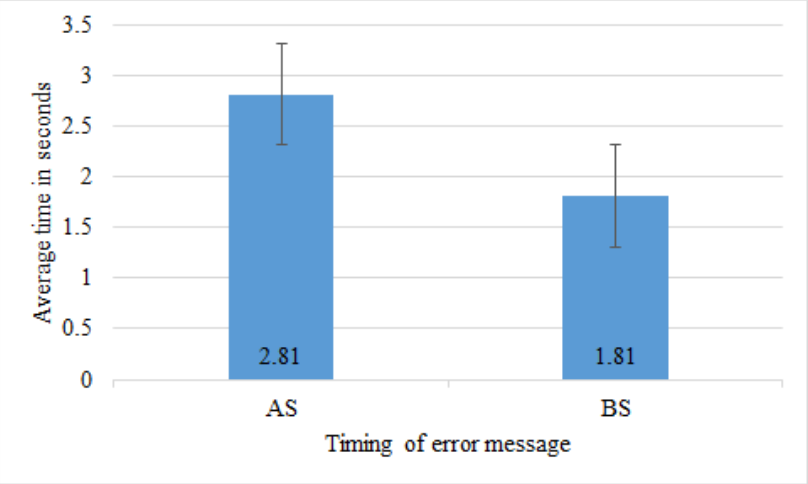

FIGURE 4.12: Average completion time of timing tested groups using small device.

TABLE 4.27: Time to complete form, Timing is tested for big screens

Method	N	M	SD
AS	33	1.7730	.5809
BS	33	1.5106	.0743
Total	66	1.6418	.5229

*Big screen. Error message Timing is tested.

Table 4.27 carries data related to the total time needed to complete the entire form. The mean and standard deviation is listed for the tested groups (AS, BS). Big device is tested. The users of AS forms have a total time of 58.51 seconds, while BS group scored a total time of 49.85 seconds. BS produced the shorter completion time. A one-way ANOVA confirmed that the two groups yielded significant differences regarding form completion time, $F(1, 64) = 4.371$, $p = 0.041$. Figure 4.13 represents the average time taken to complete the forms.

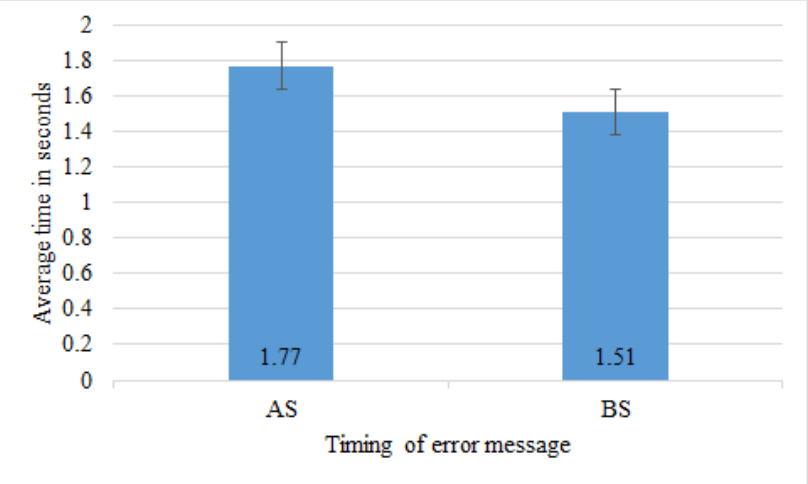

FIGURE 4.13: Average completion time of timing tested groups using big device.

4.3.2 Effectiveness

TABLE 4.28: Error Rate, location tested for small screens

Method	N	M	SD
AL	12	2.00	1.044
BL	12	2.08	.900
BT	12	1.33	.561
PU	12	2.08	.669
Total	48	1.88	.866

*Small screen. Error message location is tested.

The results provided by one-way ANOVA have confirmed that the error rate in experiment 2 for the small device where the location of error messages was tested, is not statistically different for the tested groups (AL, BL, BT, PU). F(3, 44) = 2.284, and p = 0.092. In other words, the number of the errors made by the users don't vary among the groups using the small device. Table 4.28 contains the mean and standard deviation of the tested groups and figure 4.14 shows the error rate for the location tested groups using small device.

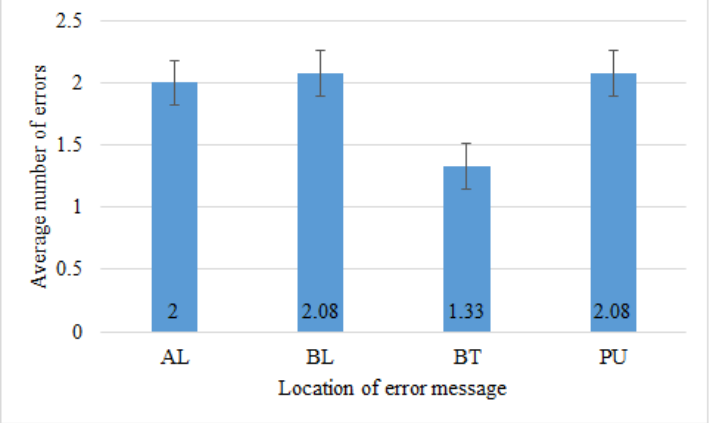

FIGURE 4.14: Error rate for location tested groups using small device.

TABLE 4.29: Error Rate, location tested for big screens

Method	N	M	SD
AL	8	1.88	.835
BL	8	1.00	.000
BT	8	1.63	.744
LB	8	1.13	.354
PU	8	1.25	.886
RB	8	1.25	.886
Total	48	1.35	.729

*Big screen. Error message location is tested.

Table 4.29 gives the mean and standard deviation of the compared groups (AL, BL, BT, PU, RB, LB). However, the results provided by one-way ANOVA have

confirmed that the error rate in experiment 2 for the big device where the location of error messages was tested, is not statistically different for the tested groups. $F(5, 42) = 1.773$, and $p = 0.139$. In other words, the number of the errors made by the users don't vary among the groups using the big device. Figure 4.15 shows the error rate for the location tested groups using big device.

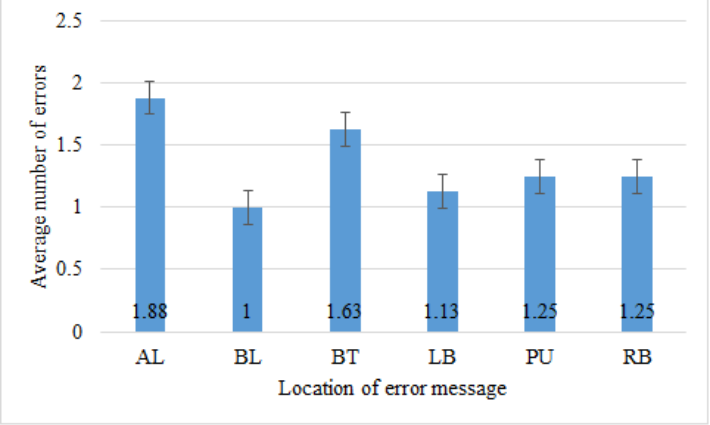

FIGURE 4.15: Error rate for location tested groups using big device.

TABLE 4.30: Error Rate, Timing tested for small screens

Method	N	M	SD
AS	12	2.00	.603
BS	12	1.33	.651
Total	24	1.67	.702

*Small screen. Error message Timing is tested.

After interpreting the results provided by one-way ANOVA, we have found that the error rate in experiment 2 where the timing of error message is tested for the small device, is significantly different among the tested groups (AS, BS). $F(1, 22) = 6.769$, $p = 0.016$. Users who have to correct the errors before submitting the form have made less errors than those who corrected the errors after submitting the form. Table 4.30 provides the mean and standard deviation of the compared groups. Figure 4.16 shows the error rate for the timing tested groups using small device.

TABLE 4.31: Error Rate, timing tested for big screens

Method	N	M	SD
AS	33	1.33	.777
BS	33	0.90	.579
Total	66	1.12	.713

*Big screen. Error message Timing is tested.

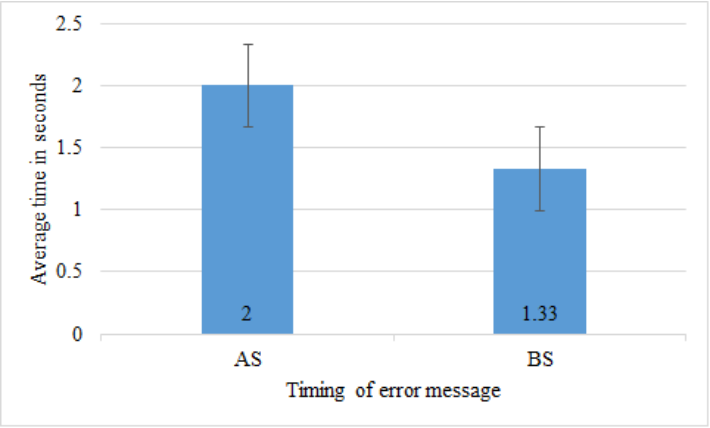

FIGURE 4.16: Error rate for timing tested groups using small device.

The results given by one-way ANOVA indicated that the error rate in experiment 2 where the timing of error message is tested for the big device, is significantly different among the tested groups (AS, BS). $F(1, 64) = 6.323$, $p = 0.014$. Users who have to correct the errors after submitting the form have made more errors than those who have corrected the errors before submitting the form. Table 4.31 provides the mean and standard deviation of the compared groups and figure 4.17 represents the error rate.

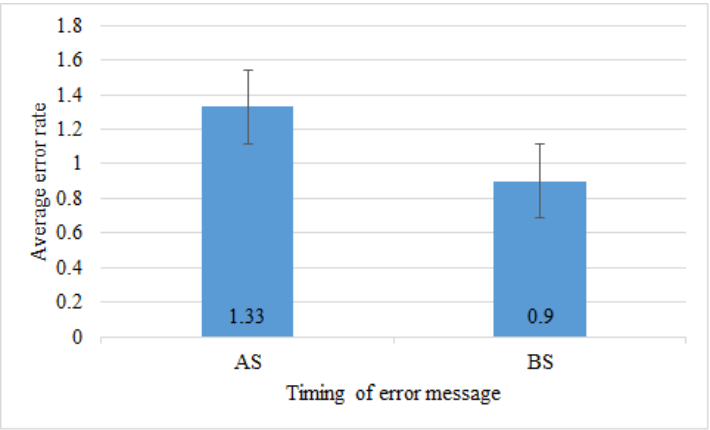

FIGURE 4.17: Error rate for timing tested groups using big device.

TABLE 4.32: Consecutive Error rate, location tested for small screens

Method	N	M	SD
AL	12	1.42	1.379
BL	12	1.17	1.403
BT	12	.58	1.240
PU	12	.33	.651
Total	48	.88	1.248

*Small screen. Error message Location is tested.

Table 4.32 provides a list of the tested groups alongside with each group computed mean and standard deviation. A one-way ANOVA test was carried to determine the significance between tested groups. The number of consecutive errors is not statistically significant between the experimented groups for the small device. F(3, 44) = 2.076, p =0.117. Therefore, we concluded that the different locations of error messages using the small screen dont affect the consecutive error rate. Figure 4.18 shows the consecutive error rate for the location tested groups using small device. Table 4.33 provides a list of the tested groups

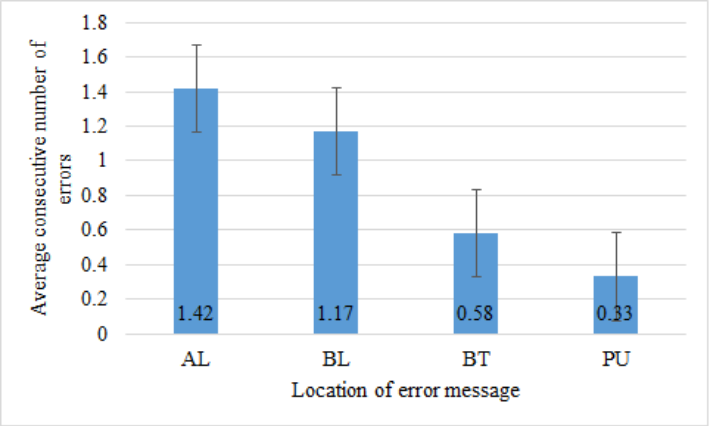

FIGURE 4.18: Consecutive Error Rate for location tested groups using small device.

TABLE 4.33: Consecutive Error rate, location tested for big screens

Method	N	M	SD
AL	8	.50	.756
BL	8	.13	.354
BT	8	1.13	.835
LB	8	.88	.991
PU	8	.63	.518
RB	8	.25	.463
Total	48	.58	.739

*Big screen. Error message Location is tested.

alongside with each group computed mean and standard deviation. A one-way ANOVA test was carried to determine the significance between tested groups. The number of consecutive errors is not statistically significant between the experimented groups for the big device. F(5, 42) = 2.0380, p =0.055. Therefore, we concluded that the different locations of error messages using the big screen dont affect the consecutive error rate. Figure 4.19 shows the consecutive error rate for the location tested groups using big device.

Table 4.34 reports the mean and standard deviation of each experimented group

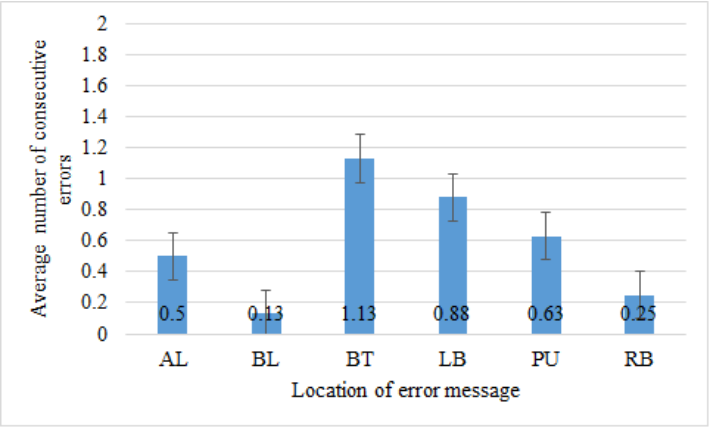

FIGURE 4.19: Consecutive Error Rate for location tested groups using big device.

TABLE 4.34: Consecutive Error rate, timing tested for small screens

Method	N	M	SD
AS	12	1.08	1.165
BS	12	.50	1.000
Total	24	.79	1.103

*Small screen. Error message Timing is tested.

(AS, BS) for the small device. There wasnt a statistically significant difference between the groups regarding consecutive error rate as it was determined by one-way ANOVA. F(1, 22) = 1.733, p = 0.202. Figure 4.20 shows the consecutive error rate for the timing tested groups using small device.

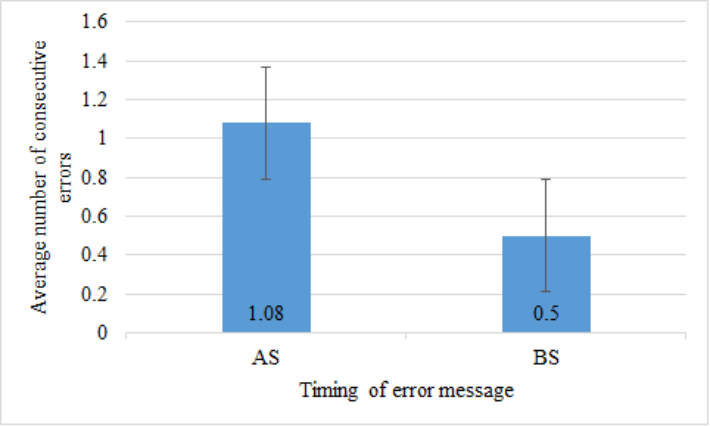

FIGURE 4.20: Consecutive Error Rate for timing tested groups using small device.

TABLE 4.35: Consecutive Error rate in experiment 2

Method	N	M	SD
AS	33	0.85	.667
BS	33	0.48	.667
Total	66	0.67	.687

*Big screen. Error message Timing is tested.

Table 4.35 reports the mean and standard deviation of each experimented group (AS, BS) for the big device. There was a statistically significant difference between the groups regarding consecutive error rate as it was determined by one-way ANOVA. $F(1, 64) = 4.902$, $p = 0.03$. Figure 4.35 represents the consecutive error rates for the tested groups.Figure 4.21 shows the consecutive error rate for the timing tested groups using big device.

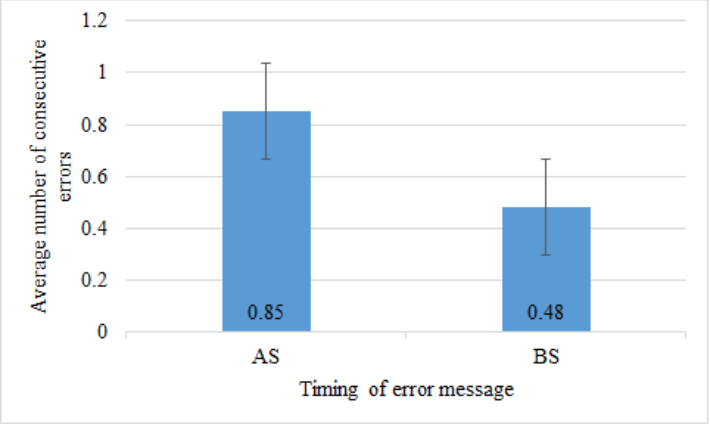

FIGURE 4.21: Consecutive Error Rate for timing tested groups using Big device.

4.3.3 Satisfaction

TABLE 4.36: Subjective ratings, small screens, location is tested

Question	Location							
	AL		BL		BT		PU	
	M	SD	M	SD	M	SD	M	SD
EL was not disturbing	2.58	1.084	2.92	.793	3.17	.577	3.17	.577
EM was helpful	3.25	.754	2.83	.937	3.00	.426	3.25	.866
EM was easy to correct	3.08	.669	3.08	.669	3.00	.603	3.58	.515

*Small screen. Location is tested.

Regarding subjective preferences and satisfaction where the location of error messages and users have tested with the small screen, the users were asked to answer the three questions that is implemented for example in table 4.36. For question 1, EL was not annoying, as come to those who tested the forms with errors being displayed above the label (AL), 2 (6.25%) answered with strongly agree, 6 (50%) answered with agree, 1 (6.25%) answered with disagree and 3 (25% with strongly disagree.

For BL group, 2 (16.67%) have answered with strongly agree, 8 (66.67%) with agree, 1 (6.25%) with disagree, and 1 (6.25%) with strongly disagree. BT group,

3 (25%) have answered with strongly agree, 8 (66.67%) with agree, 1 (8.33%) with disagree, and none answered with strongly disagree. And finally, for PU participants, the results was as for BT group. A one-way ANOVA has confirmed that there is no significant differences between the tested groups (AL, BL, BT, and PU), $F(3, 44) = 1.485$, $p = 0.232$.

The second question: EM was helpful, 14 (29.67%) answered strongly agree, 26 (54.17%) agree, 6 (12.5%) disagree, and 2 (4.17%) strongly disagree, which suggests that the involved users found the error messages helpful. However, ANOVA test has shown no significant differences between the tested groups, $F(3, 44) = 0.841$, $p = 0.479$. The third question: the error is easy to correct, 15 (31.25) answered with strongly agree, 27 (56.25%) have answered with agree, 6 (12.5%) have answered with disagree and none of the participants have answered with strongly disagree. The previous numbers indicate that most of the participants found that the error easy to correct. A one-way ANOVA has shown that there is no significant differences among the groups, $F(5, 42) = 1.\ 211$, $p = 0.321$. Figure 4.22 shows the subjective rating for the location tested groups using small device. Regarding subjective preferences and satisfaction where the location of

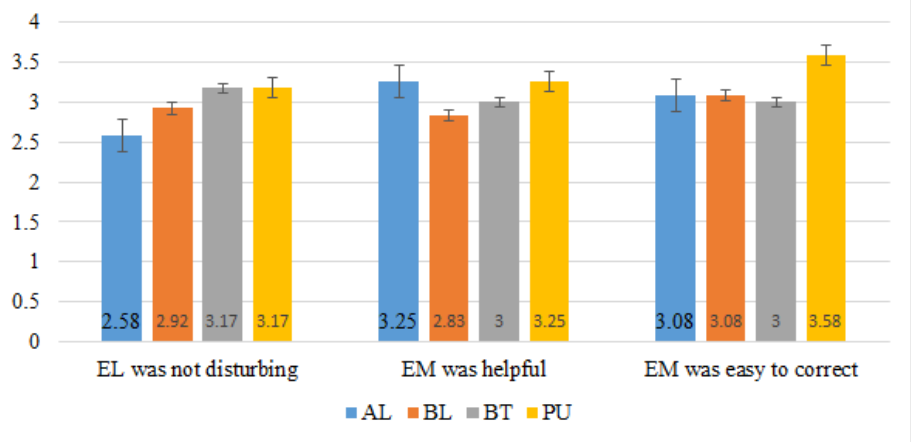

FIGURE 4.22: Subjective rating of location of error messages using small device.

error messages and users have tested with the big screen, the users were asked to answer the three questions that is implemented for example in table 33. For question 1, EL was not annoying, as come to those who tested the forms with errors being displayed above the label (AL), 3 (37.5%) answered with disagree, 4 (50%) answered with agree, 2 (25%) answered with strongly agree.

For BL group, 1 (12.5%) have answered with disagree, 6 (75%) with agree, 1

TABLE 4.37: Subjective ratings, location is tested, big screens

Question	Location					
	AL	BL	BT	LB	PU	RB
	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
EL was not disturbing	2.88 (.835)	3.00 (.535)	2.75 (.886)	.340 (.516)	.250 (.535)	.317 (.408)
EM was helpful	3.25 (.707)	3.25 (.463)	3.00 (1.069)	3.10 (.316)	2.75 (.463)	3.33 (.816)
EM was easy to correct	3.25 (.707)	3.13 (.354)	3.00 (.756)	3.10 (.316)	2.63 (.518)	2.83 (.7530)

*Big screen. Location is tested.

(12.5%) with strongly agree. BT group, 4 (50%) have answered with disagree, 2 (25%) with agree, 2 (25%) answered with strongly agree. For PU participants, 4 (50%) have answered with disagree, and 4 (50%) with agree. the results was as for BT group. For RB group, 5 (62.5%) responded with agree, and 3 (37.5%) responded with strongly agree. Finally, for LB group, 6 (75%) have responded with agree and 2 (25%) answered with strongly agree. A one-way ANOVA has confirmed that there is no significant differences between the tested groups (AL, BL, BT, PU, RB, and LB), $F(5, 42) = 2.055$, $p = 0.090$.

The second question: EM was helpful, 1 (2.08%) answered strongly disagree, 5 (10.42%) disagree, 30 (62.5%) agree, and 12 (25%) strongly agree, which suggests that the involved users found the error messages in general helpful. However, ANOVA test has shown no significant differences between the tested groups, $F(5, 42) = 0.783$, $p = 0.568$. The third question: the error is easy to correct, 8 (16.67%) disagree, 32 (66.67%) agree, and 8 (16.67%) strongly agree. The previous numbers indicate that most of the participants found that the error easy to correct. A one-way ANOVA has shown that there is no significant differences among the groups, $F(5, 42) = 1.\ 211$, $p = 0.321$. Figure 4.23 shows the subjective rating for the location tested groups using big device. Participants have

TABLE 4.38: Subjective ratings, small screens, timing is tested

Timing	AS		BS	
Question	M	SD	M	SD
ET wan not disturbing	3.08	.669	3.17	.389
EM was helpful	3.00	.426	3.00	.426
EM was easy to correct	3.00	.603	3.00	.603

*Small screen. Timing is tested.

tested the timing (AS, BS) with the small screen, have answered the questions

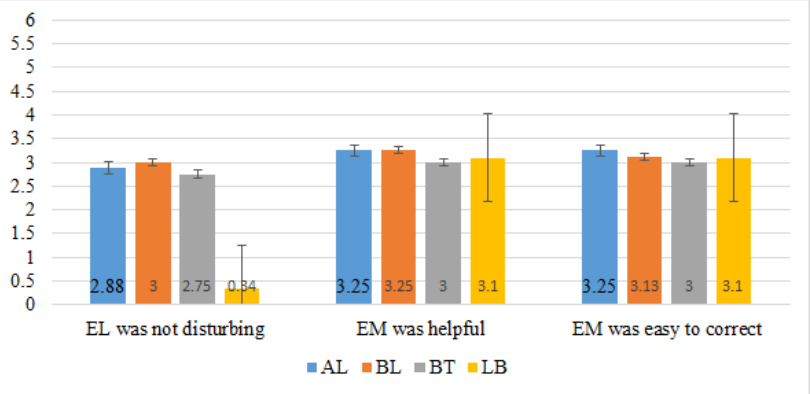

FIGURE 4.23: Subjective rating of location of error messages using big device.

involved in table 4.38. The responses on question 1 was as the following: 2 (16.67%) answered disagree, 7 (58.33%) answered agree, and 3 (25%) responded with strongly agree. The numbers implies that most users that tested AS forms found the timing of the error message not disturbing at all. The other participants engaged in BS group responses was as the following: 10 (83.33%) agree, and 2 (16.67%) strongly agree. A one-way ANOVA revealed that no significant differences that exist between the two tested groups (AS, BS), $F(1, 22) = 0.139$, $p = 0.713$.

The responses on question 2 was as the following: 2 (8.33%) answered disagree, and 20 (83.33%) responded with agree, and 2 (8.33%) strongly agree. The numbers implies that most users found the error messages helpful. A one-way ANOVA revealed that no significant differences that exist between the two tested groups (AS, BS), $F(1, 22) = 0.000$, $p = 1000$.

The responses on question 3 was as the following: 3 (12.5%) answered disagree, and 16 (66.67%) responded with agree, and 4 (16.67%) strongly agree. The numbers implies that most users found the error messages easy to correct. A one-way ANOVA revealed that no significant differences that exist between the two tested groups (AS, BS), $F(1, 22) = 0.000$, $p = 1.000$. Figure 4.24 shows the subjective rating for the timing tested groups using small device. Participants

TABLE 4.39: Subjective ratings, big screens, timing tested groups

Timing	AS		BS	
Question	M	SD	M	SD
ET wan not disturbing	2.94	.556	3.03	.637
EM was helpful	3.06	.556	2.94	.123
EM was easy to correct	3.00	.559	3.00	.107

*Big screen. Timing is tested.

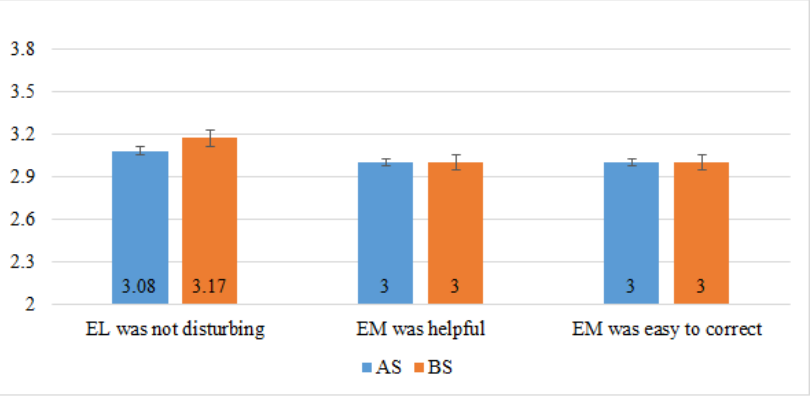

FIGURE 4.24: Subjective rating of timing of EM using small device.

have tested the timing (AS, BS) with the big screen, have answered the questions involved in table 4.39. The responses on question 1 was as the following: 6 (18.18%) answered disagree, 23 (69.69%) answered agree, and 4 (12.12%) responded with strongly agree. The numbers implies that most users that tested AS forms found the timing of the error message not disturbing at all. The other participants engaged in BS group responses was as the following: 5 (15.15%) disagree, 20 (60.61%) agree, and 8 (24.24%) strongly agree. A one-way ANOVA revealed that no significant differences that exist between the two tested groups (AS, BS), $F(1, 64) = 0.382$, $p = .539$.

The responses to question 2 was as the following: 2 (3.03%) answered strongly disagree, 4 (6.06%) answered disagree, and 48 (72.73%) responded with agree, and 10 (13.63%) strongly agree. The numbers implies that most users found the error messages helpful. A one-way ANOVA revealed that no significant differences that exist between the two tested groups (AS, BS), $F(1, 64) = 0.226$, $p = 0.642$.

The responses to question 3 was as the following: 10 (15.15%) responded with disagree, 43 (65.15%) agree, and 10 (15.15%) strongly agree. The numbers implies that most users found the error messages easy to correct. A one-way ANOVA revealed that no significant differences that exist between the two tested groups (AS, BS), $F(1, 64) = 0.602$, $p = 0.441$. Figure 4.25 shows the results of the subjective ratings of tested groups. As to the preferred location according to the participants who have tested small device: 11 (22.91%) reported that they preferred the message to be placed above the label (AL), 15 (31.25%) answered they consider BL as the best location, 13 (27%) voted for BT location, and 9 (18.75%) have chosen PU as the best location. The results are shown by figure

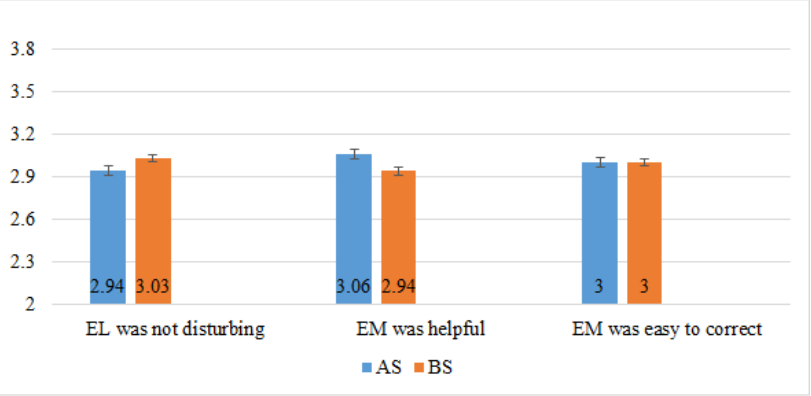

FIGURE 4.25: Subjective rating of timing of error messages using big device.

4.26. However, the choices of the participants who have tested the big screen: 7

FIGURE 4.26: Best location according to small device users.

(14.58%) preferred the above label location (AL), 11 (22.91%) considered BL as the best location, 15 (31.25%) participants have preferred BT, 3 (6.25%) voted for PU, 6 (12.5%) voted for LB, and 6 (12.5%) voted for RB as it is shown by figure 4.27. As for the timing of error message (AS, or BS), the participants that used small device: 9 (37.5%) have chosen after submission as the best timing for the error message to appear, while 15 (62.5%) voted for BS. Results are presented by figure 4.28. As for the timing of error message (AS, or BS), the participants that used big device: 21 (31.8%) have chosen after submission as the best timing for the error message to appear, while 45 (68.2%) voted for BS. Figure 4.29 presents the results. The above numbers suggest that in general that for both devices the users tend to prefer error messages to appear immediately while filling each field (BS) over AS.

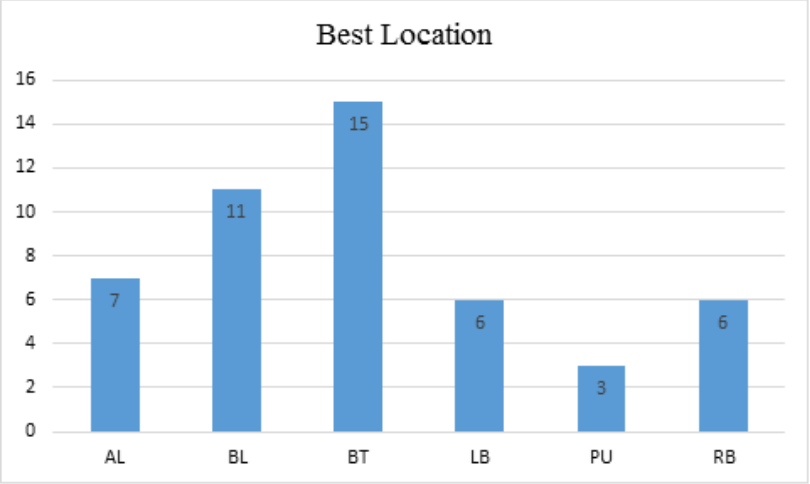

FIGURE 4.27: Best location according to big device users.

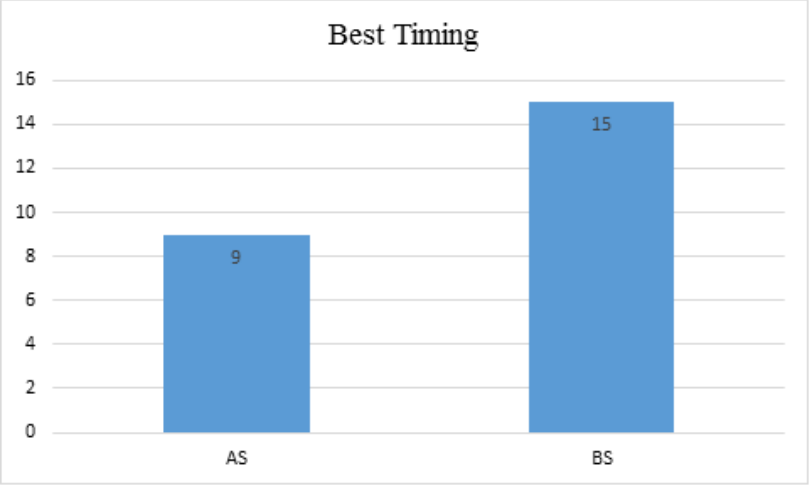

FIGURE 4.28: Best Timing according to small device users.

4.4 Effects of graphics on form filling

4.4.1 Efficiency

TABLE 4.40: Total time to correct the errors

Method	N	M	SD
BT	33	.4797	.4411
HB	33	.2782	.2337
Total	66	.3789	.0449

*Big screen. Use of Graphics is tested.

The mean and standard deviation regarding the time taken to complete the entire form are provided by Table 4.40. A one-way ANOVA test has revealed that there is a statically significant differences between the two groups (BT, HB). $F(1, 64) = 5.379$, $p = 0.024$. BT errors took longer time to be corrected than HB errors. Figure 4.30 represents the average time needed to correct the errors made by the

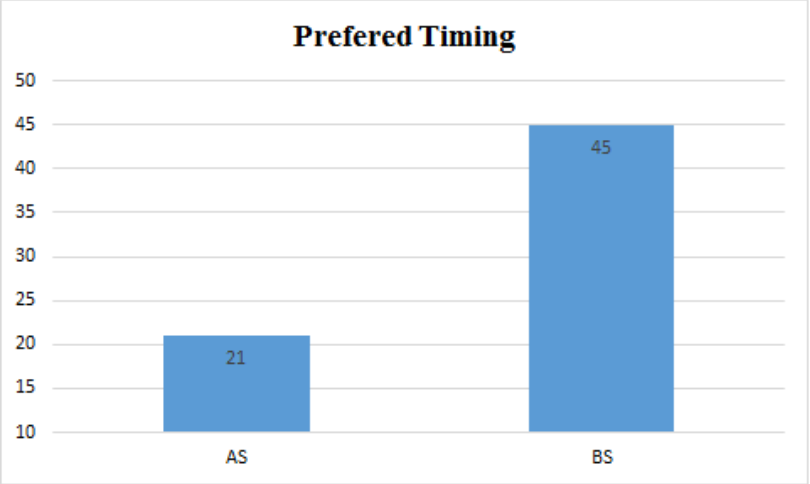

FIGURE 4.29: Best Timing according to big device users.

two groups.

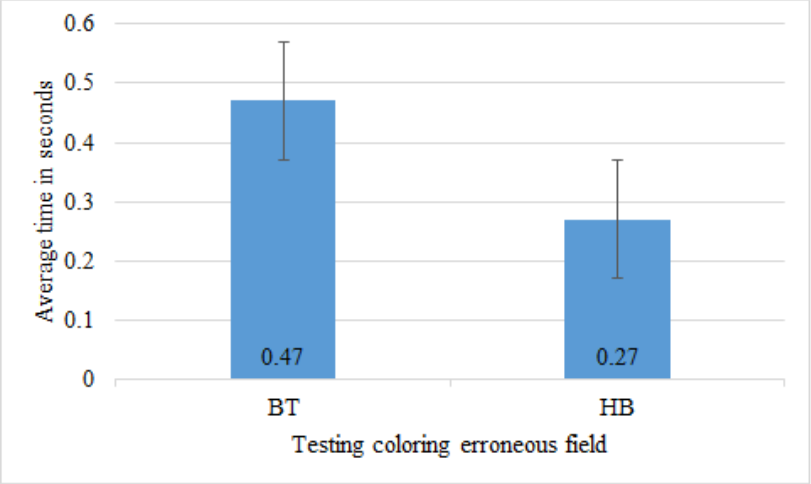

FIGURE 4.30: Time to correct all errors, experiment 3.

TABLE 4.41: Time to complete the form, experiment 3

Method	N	M	SD	
BT	33	2.00	.57908	*
HB	33	1.6042	.07553	
Total	66	1.8021	.06714	

Big screen. Use of Graphics is tested.

The mean and standard deviation regarding the time taken to complete the entire form are provided by Table 4.41. A one-way ANOVA test has spotted that there is a statically significant difference between the two groups (BT, HB). $F(1, 64) = 9.871$, $p = 0.003$. BT forms took longer time to be completed than HB forms. Figure 4.31 shows the average completion time for the tested groups.

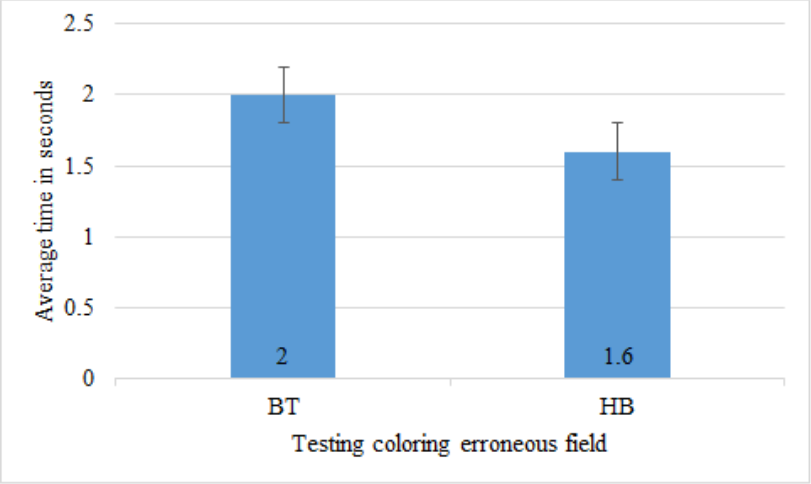

FIGURE 4.31: Average Form Completion Time, experiment 3.

4.4.2 Effectiveness

TABLE 4.42: Error rate in experiment 3

Method	N	M	SD
BT	33	1.36	0.783
HB	33	1.36	0.549
Total	66	1.36	0.671

*Big screen. Use of Graphics is tested.

In experiment 3 we have tested the forms against errors. The error rate wasnt significantly different between the two groups (BT, HB) as it was confirmed by one-way ANOVA. The mean and standard deviation are available in Table 4.42. $F(1, 64) = 0.000$, $p = 1.000$.Figure 4.32 shows the error rates regarding the tested groups. The mean and standard deviation regarding the consecutive

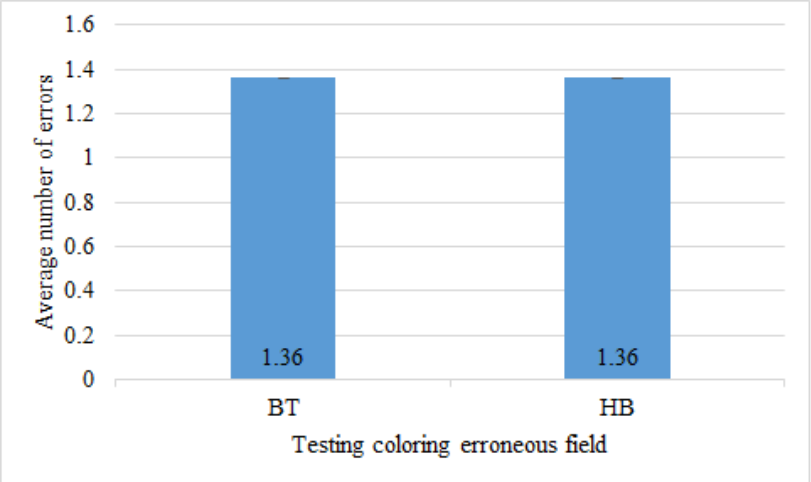

FIGURE 4.32: Error Rates in experiment 3.

error rate are provided by Table 4.43. A one-way ANOVA test has confirmed

TABLE 4.43: Consecutive Error rate in experiment 3

Method	N	M	SD
BT	33	0.76	.751
HB	33	0.36	.603
Total	66	0.56	.704

*Big screen. Use of Graphics is tested.

that consecutive error rate is significantly different between the two groups (BT, HB). $F(1, 64) = 5.518$, $p = 0.022$. Users who participated in filling the forms that involved highlighting the erroneous input field (HB) have committed less errors than those engaged in the other group that didnt involve the use of graphics (BT). Figure 4.33 shows the consecutive error rates of the tested groups.

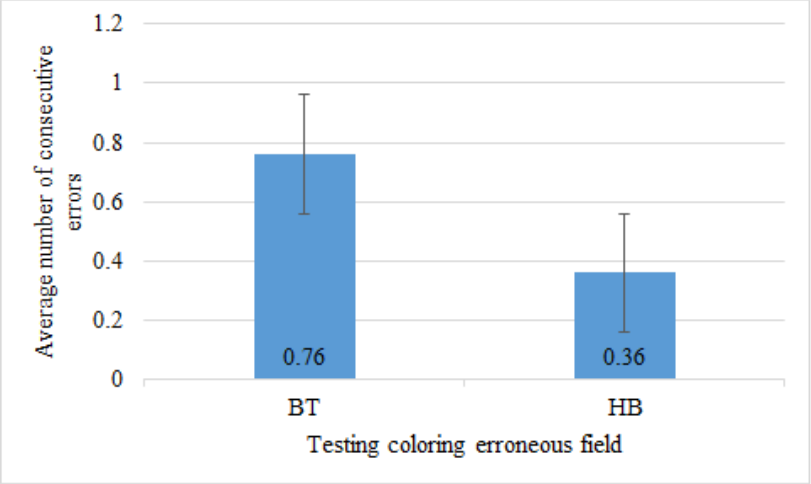

FIGURE 4.33: Consecutive error rate, experiment 3.

4.4.3 Satisfaction

In experiment 3 where the use of graphics was tested, 66 participants were involved. 60 (90.91%) have supported the use of graphics while only 6 (9.09%) provided a negative answer regarding the use of graphics.

The answers for whether highlighting the erroneous field have lead the participant to notice the error faster than the other case (providing the error message in red only): 30 (90.91%) participants have admitted that highlighting the erroneous input field indeed have led them to notice the error location quickly, and 3 (9.09%) disagreed regarding the matter. On the other hand, in BT category, 28 (84.85%) participants have said that if the box was colored, it will lead them

to locate the error faster, and only 5 (15.15%) responded with a no.

Chapter 5

Discussion and Conclusion

5.1 Discussion

5.1.1 Screen size effect on error rate

Smartphones come with different capabilities and sizes. We wanted to examine if there is a relationship between the size of the screen and the number of errors the user make during form filling since it was never been investigated before by other researchers. Our experiment has proven the existence of such relationship. The number of errors made by the users of the big screen is much less than the number of errors made by the users of the small screen. The feedback that we got from the participants supported our findings. The users' feedback included statements regarding screen size and its impact on the number of errors made by them. The majority of the small screen participants expressed their frustration that resulted from having to fill a form with a small device. We hypothetically assumed that the small real estate of the device requires more focus and users' attention to read the instructions provided, as well as the presented error. In addition, we have assumed that the keyboards' small buttons may generate more typos.

The feedback provided by the participants supported our assumptions. The participants expressed their frustration regarding the issues that the small screen generates. They have said that device small screen has lead them to create more errors, some attributed that to the size of the buttons which pushed them to

hit some buttons accidantly. Others attributed that to the small font size which required more attention while reading.
Once again the feedback provided our assumptions regarding the screen size and its effect on form filling. We do encourage researchers to conduct further investigation to come up with solutions that might reduce the issues generated by small screens.

5.1.2 Testing location and timing of error messages

The detailed literature review that was carried for this dissertation has demonstrated that forms error messages can be placed in different locations. Popular locations include above the label, below text boxes, placeholders, above the forms, dialogue boxes and others. We have chosen to evaluate a set of them through empirical testing. The testes was carried for two different screen sizes. For the small screen, we have picked the following locations: above the label, below label, pop up, and below text boxes. We have evaluated the locations and timing of error messages regarding efficiency, effectiveness, and satisfaction. Efficiency was measured by the time taken to complete the forms, time taken to correct all the errors and time taken to correct each type of errors separately.
The findings of the small device regarding the location of error messages with respect to the completion time, error correction time and consecutive error rates indicated that when error messages appeared to the users above the label scored the worst results among the tested groups. Perhabs we can attribute that to the fact that users are not used to see error messages to be displayed above the label. Our findings regarding error rates has indicated that pop up message has achieved the worst results (PU users have commited more errors than the rest of the groups). Our guessing is that the reason PU users scored more errors refer to the fact that error messages of BL, BT, and AL groups are located near to the erroneous field, so users dont have to look far away to recheck the content of the presented error messages. In addition, the error message dont disappear like the messages in PU group, if the user want to reread the message in PU groups he need to hit the submit button again in order to see the message which implies some additional time to the filling process.
However, the analysis of the collected data shown that the differences between

the means of tested groups are not statistically significant. Which in turn implies that the time to complete the forms is close for all of the tested groups. Our findings matches to some level the findings of [7] regarding error messages in web forms. However, our testing of the pop up messages involved the presentation of error messages after clicking the submit button, while their work involved immediate feedback.

The previous discussion applies to the big device groups as well, except that the group that scored the shortest time is the BL group. In our openion and according to the feedback that we got from users in rertospective think aloud placing error messages below label scores the best results because users will immediatly recognize that the error message belong to that field, because it is placed between the labe and its text box.

5.1.3 Testing the use of graphics

The analysis of the data collected in experiment 3 have shown a statistically significant difference regarding form completion time. The HB group scored the shortest time needed to complete the forms.

The evaluation of the timing groups yielded much better results. For the small device groups, the analysis indicated the existence of a significant difference between the tested groups. BS group scored the shortest time among the two groups. Our findings support the guidelines offered by the study that was carried by [5] On the other hand, the analysis of big screen timing groups AS scored shorter time than BS. However, the analysis has shown no significant results regarding timing groups of big screen. It is worth to mention that BS and AS form completion time was relatively close. That matches to some level the findings of [7] regarding presenting the errors all at one after submission for web forms.

The results regarding the time taken to correct all errors made by each group indicated the following: both the small device and big device with respect to the location of the tested groups, the analysis have shown that there is statistically significant differences between the tested groups, which contradict the findings of [21] that declared that error correction time is not statistically significant between the different error messages location for web forms. BT has scored the shortest time among the small screen groups while PU scored the shortest time

among the big screen groups. However, the analysis of the timing groups have revealed no significant differences between the groups. For both devices, BS once again scored the shortest time needed to correct the errors.

The analysis of the data collected in experiment 3 have shown a statistically significant difference regarding time required to correct the errors. The HB group scored the shortest time needed to correct the errors.

Time needed to correct each type of errors separately yielded no significance between the tested groups for both test categories: location and timing. However, the password field scored the longest time that the user needed to correct it. Our explanation to this outcome is due to the several obstacles imposed on the password field. The findings contradicts [37]. They have found no differences between tested groups regarding time needed to correct password errors.

Effectiveness was measured by error rate and consecutive error rate. As for error rate which is the first time errors, the small and the big device with respect to the location of error messages were tested. The analysis have shown no significant differences among the tested groups for both devices. The group with the fewest number of errors is BT regarding the small device while BL is the group with fewest number of errors regarding the big device. On the other hand, results regarding the timing of error messages, the small device data analysis have shown a significant difference between the tested groups. That contradicts the results of [7] regarding the number of first time errors for web forms. BS scored less number of errors. No statistical significant differences are found between the big device tested groups, in fact, both groups have scored the same number of errors.

Consecutive error rates for both devices dont vary greatly between the groups as it was suggested by the analysis. The group that scored the fewest number of consecutive errors is PU for the small device, while BL scored the fewest number of consecutive errors among the big device groups. BS scored the fewest errors regarding both devices.

The analysis of the data collected in experiment 3 have shown a statistically significant difference regarding error rate, time to correct errors and time to complete the forms. The HB group scored the fewest number of errors, less time taken to correct the errors and to complete the forms. That leads to encourage the use of graphics in form design.

Satisfaction was measured by subjective rating and feedback given by the participants. The small device users voted for BL as the best location, while users of big screen voted for BT as the best place to present the error messages. Both of the user categories voted for PU as their least favorite. As we discussed before, PU involves presenting the error message as a pop up box that appears for some time then disappears quickly which requires to show the message again if the user wants recheck the content of the error message. In addition, the message appear at the end of the screen, which is quite the distance between the first few boxes and the error messages.

For both devices, the users picked BS as the best timing to show the error message. Immediate feedback decreases the number of submission the user has to make. It also display the error messages one by one which wont lead the user to get confused as if the errors were displayed all at once. Especially if the user is confronting a large number of errors.

5.2 Conclusion

The proliferation of the idea of mobile devices has led the cyber age to a new era. This breakthrough altered research efforts and directed them to new routes. Practitioner now need to come up with new design trends to accommodate the diversity associated with these devices. Mobile devices such as tablets, smartphones, and PDAs properties are different. Our work concentrates on smartphones. Smartphones diverse sizes encouraged us to choose two devices for our experiments, each one of them associated with a different screen real estate.

Three experiments were conducted to answer three main questions. Experiment one addressed the relationship between screen size and number of error made by the users using form filling. Our findings demonstrated that the small screen generated more errors than the big screen. This signifies an important issue than must be considered during the design process. Designers must give small screens a considerable amount of efforts and to start considering to involve shorter forms for screens with small real estate and to provide assistance as much as possible for users during form filling. Experiment two answered the question, regarding the best location and timing of error messages. Small device: We have found

that BT achieved the best results regarding error rate and time to correct errors, while PU achieved superior results regarding the number of consecutive error rate. We consider BT as the best among the location categories, since PU has close results to BT. Big device: BL has achieved the best results regarding all measures. As for the best timing BS scored superior results in comparison with AS.

Experiment three addressed the question regarding graphics. The results and their analysis have shown that the use of graphics signified the statistical differences between the tested groups. These findings encourages the use of graphics to facilitate error messages presentation. We found no evidence to prove the effect of the culture on error messages representation. We attribute that to fact that we only involved text boxes and we have allowed the use of both languages (English and Arabic) to enter the data. We received zero complaints regarding the used languages from the users. They have shifted back and forth between the languages whenever it was required.

We provide the following set of mobile forms design guidelines based on the findings of this dissertation:

1. Small devices best error location is below the text field. Pop up messages are recommended to present errors one by one and not all at once.

2. Big screens best error location is below the label.

3. Best timing to present the error message is before submission, where immediate input validation and feedback is given to the user.

4. Presenting error messages above the label is not favorable and it is not recommended.

5. Although, big screens have enough real estate to present errors beside the input boxes it is not recommended.

6. Facilitate error presentation with graphics. It captures the attention of the user to the location of the error faster than the traditional way of error presentation.

Chapter 6

Limitations and Future Work

This section covers the set of limitations imposed on our work that we are hoping to overcome in the future research. In addition, this section includes future work. The kind of work we are hoping to accomplish in the future based on the findings of this research.

6.1 Limitations

There are quite number of limitations that have to be addressed. First, there are several number of possible locations where the error message might appear, we have addressed only a portion of them. We have neglected for example, testing error messages where they appear beside the label or inside the input field (also known as placeholders). Second, we have tested two possibilities when it comes to the timing of the error messages (after submit, and before submit). There are other popular methods that is used to present error messages, for example, in our experiments, we have designed the forms in such a way that when the user clicks the submit button, he or she will receive all the errors at once. Other methods that we have neglected is to receive the errors one by one. Third, we have tested the timing of error messages in isolation, without considering the possible effects that the location (which is diverse as it was discussed above) of the error message might has on users' performance.

Forth, the number of participants involved in experiment three was relatively small. To generalize the results we would involve more participants. Fifth, our

tests were targeting computer major students. Students that surly have some experience regarding computer, internet, forms, and their components. In our tests we have neglected novice users. This segment of users might yield different types of results (for instance, longer form completion time).
Sixth, our samples were selected from the same university that is located at Palestine which implies that the groups of people we have involved in our tests belong to the same culture. As a consequence, the findings of this study cant be applied to other cultures. However, our investigations were directed towards Arabic forms in the first place, so the applicability to other cultures doesnt matter at this point. Seventh, the tested forms in experiment two and three were short, only few input fields were involved. Long forms might lead to different sorts of findings.
Eighth, we have involved text input boxes, other type of input boxes include calendars, or radio buttons might affect the results. Finally, the users provided their ratings of the best location and timing without truly having the experience with the different locations and timings of error messages. So, may be if they tried them all, it may provide a better comparative process, and perhaps they change their mind about their answers regarding their ratings of location and timing of error messages.

6.2 Future Work

Our future work will include overcoming the limitations imposed on this work. We will increase the number of participants, and we will surely involve novice users to investigate the impact they may bring on the findings proposed by this dissertation. In addition, we will investigate long forms or a long form that is decomposed into multiple smaller forms and the navigation would be through the use of next and previous buttons. In our future work, we seek to examine the presentation of error messages (location and timing at the same time), and test showing error messages after submission, one by one. Future work may also include exploring other types of input fields such as radio buttons, drop down lists, and others. In order to check if the results given by the forms that only contains text boxes still hold to the forms that contains multiple types of

input fields. Further research may include the evaluation of other locations of error messages. Finally, we look forward to deploy an eye tracking device in our experiments. These devices gain the researchers a better insight and knowledge regarding the human behavior, like what is first thing that captures the human eye, did the participant notice the error message immediately and others. Such kind of information could be helpful, especially for experiment three, since our main goal is to examine the effect of graphics and whether they lead to achieve higher affordance (error message noticed faster). Error recognition time can be accurately captured and extracted with the availability of eye tracking tools.

Bibliography

[1] Al-Saleh, M., Al-Wabil, A., Al-Attas, E., Al-Abdulkarim, A., Chaurasia, M., and Alfaifi, R. Inline immediate feedback in arabic web forms: An eye tracking study of transactional tasks. *International Conference on Innovations in Information Technology (IIT), IEEE* (2012), 333–338.

[2] Albert, W., and Tullis, T. Measuring the user experience: collecting, analyzing, and presenting usability metrics. *Newnes* (2013).

[3] Ali, A., Alrasheedi, M., Ouda, A., and Capretz, L. A study of the interface usability issues of mobile learning applications for smart phones from the users perspective. *arXiv preprint arXiv:1501.01875* (2015).

[4] Alton, N., Rinn, C., Summers, K., and Straub, K. Using eye-tracking and form completion data to optimize form instructions. *International Professional Communication Conference (IPCC), IEEE* (2014), 1–7.

[5] AnswerLab, and Google. Principles of mobile site design: Delight users and drive conversions. *https://storage.googleapis.com/think-v2-emea/docs/article/multi-screen-moblie-whitepaper_research-studies_1.pdf* (2014).

[6] Avoka, T. Responsive forms: 8 best practices in responsive design principles for digital forms. *http://www.avoka.com/* (2016).

[7] Bargas-Avila, J., Oberholzer, G., Schmutz, P., de Vito, M., and Opwis, K. Usable error message presentation in the world wide web: Do not show errors right away. *Interacting with Computers, Elsevier 19*, 3 (2007), 330–341.

[8] Bevan, N. Extending quality in use to provide a framework for usability measurement. *Human Centered Design, Springer* (2009), 13–22.

[9] Boiano, S., Bowen, J., and Gaia, G. Usability, design and content issues of mobile apps for cultural heritage promotion: The malta culture guide experience. *arXiv preprint arXiv:1207.3422* (2012).

[10] Bojko, A., and Schumacher, R. Eye tracking and usability testing in form layout evaluation. *Proceedings of the 38th International Symposium of Business Forms Management Association* (2008).

[11] Budiu, R., and Nielsen, J. Usability of ipad apps and websites. *Fremont, CA: Nielsen Norman Group* (2011).

[12] Buskirk, T., and Andrus, C. Smart surveys for smart phones: Exploring various approaches for conducting online mobile surveys via smartphones. *The premier e-journal resource for the public opinion and survey research community* (2012).

[13] Chae, M., and Kim, J. Do size and structure matter to mobile users? an empirical study of the effects of screen size, information structure, and task complexity on user activities with standard web phones. *Behaviour and information technology 23*, 3 (2004), 165–181.

[14] Chan, S., Fang, X., Brzezinski, J., Zhou, Y., Xu, S., and Lam, J. Usability for mobile commerce across multiple form factors. *J. Electron. Commerce Res. 3*, 3 (2002), 187–199.

[15] Charness, G., Gneezy, U., and Kuhn, M. Experimental methods: Between-subject and within-subject design. *Journal of Economic Behavior and Organization, Elsevier 81*, 1 (2012), 1–8.

[16] Duh, H., Tan, G., and Chen, V. Usability evaluation for mobile device: a comparison of laboratory and field tests. *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services, ACM* (2006), 181–186.

[17] Garvin, K., Fernanda, G., Raissa, R., and Larissa, S. Design of small interfaces. *B.S. thesis, Universitat Politecnica de Catalunya* (2013).

[18] Gunduz, F., and Pathan, A.-S. On the key factors of usability in small-sized mobile touch-screen application. *Int. J. Multimed. Ubiquitous Eng 8*, 3 (2013), 115–138.

[19] Gutwin, C., and Fedak, C. Interacting with big interfaces on small screens: a comparison of fisheye, zoom, and panning techniques. *Proceedings of Graphics Interface 2004, Canadian Human-Computer Communications Society* (2004), 145–152.

[20] Harms, J., Kratky, M., Wimmer, C., Kappel, K., and Grechenig, T. Navigation in long forms on smartphones: Scrolling worse than tabs, menus, and collapsible fieldsets. *Human-Computer Interaction, Springer* (2015), 333–340.

[21] Inal, Y. User-friendly locations of error messages in web forms: An eye tracking study. *Journal of Eye Movement Research 9*, 5 (2016).

[22] Jarrett, C., and Gaffney, G. Forms that work: Designing web forms for usability. *Morgan Kaufmann* (2009).

[23] Ji, Y., Park, J., Lee, C., and Yun, M. A usability checklist for the usability evaluation of mobile phone user interface. *International journal of human-computer interaction, Taylor and Francis 20*, 3 (2006), 207–231.

[24] Jones, M., Marsden, G., Mohd-Nasir, N., Boone, K., and Buchanan, G. Improving web interaction on small displays. *Computer Networks, Elsevier 31*, 11 (1999), 1129–1137.

[25] Juliet, R. Selecting the right method. *Ux Research* (2015).

[26] Klockar, T., Carr, D., Hedman, A., Johansson, T., and Bengtsson, F. Usability of mobile phones. *International Symposium on Human Factors in Telecommunications* (2003), 197–204.

[27] KOHL, J. Test mobile applications with i sliced up fun! , (2010).

[28] Krug, S. Don't make me think: A common sense approach to web usability. *Pearson Education, India* (2005).

[29] Kumar, N. Usability for the new pc: Mobile devices. *www.semanticscholar.org* (2013).

[30] Lazar, J., and Huang, Y. Designing improved error messages for web browsers. *Human factors and Web development* (2003), 167–182.

[31] Linderman, M., and Fried, J. Defensive design for the web: How to improve error messages, help, forms, and other crisis points. *New Riders Publishing Thousand Oaks, CA, USA* (2004).

[32] Lobo, D., Kaskaloglu, K., Kim, C., and Herbert, S. Web usability guidelines for smartphones: a synergic approach. *International journal of information and electronics engineering, IACSIT Press 1*, 1 (2011), 33.

[33] Madrigal, D., and McClain, B. Strengths and weaknesses of quantitative and qualitative research. *UX matters* (2012).

[34] Maniar, N., Bennett, E., Hand, S., and Allan, G. The effect of mobile phone screen size on video based learning. *JSW 3*, 4 (2008), 51–61.

[35] Marcial, L. A comparison of screen size and interaction technique: Examining execution times on the smartphone, tablet and traditional desktop computer. *University of North Carolina at Chapel Hill* (2014).

[36] Matera, M., Rizzo, F., and Carughi, G. Web usability: Principles and evaluation methods. *Web engineering, Springer* (2006), 143–180.

[37] Mockovak, W. Comparing the effectiveness of alternative approaches for displaying edit-error messages in web forms. *Bureau of Labor Statistics, Statistical Survey Papers* (2005).

[38] Nielsen, J., and Budiu, R. Mobile usability. *New Riders, Berkeley* (2013).

[39] Nilsson, E. Design guidelines for mobile applications. *SINTEF ICT, Jun* (2008).

[40] Page, T. Usability of text input interfaces in smartphones. *Journal of Design Research, Inderscience Publishers Ltd 11*, 1 (2013), 39–56.

[41] Paterno, F., Schiavone, A., and Pitardi, P. Timelines for mobile web usability evaluation. *Proceedings of the International Working Conference on Advanced Visual Interfaces, ACM* (2016), 88–91.

[42] Pendell, K., and Bowman, M. Usability study of a library's mobile website: an example from portland state university. *Information Technology and Libraries (Online), American Library Association 31*, 2 (2012), 45.

[43] Raptis, D., Tselios, N., Patras, R., Kjeldskov, J., and Skov, M. Does size matter?: investigating the impact of mobile phone screen size on users' perceived usability, effectiveness and efficiency. *Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services, ACM* (2013), 127–136.

[44] Recabarren, M., and Nussbaum, M. Exploring the feasibility of web form adaptation to users cultural dimension scores. *User Modeling and User-Adapted Interaction, Springer 20*, 1 (2010), 87–108.

[45] Rohrer, C. When to use, which user-experience research methods. *http://www.nngroup.com/articles/which-ux-research-methods* (2014).

[46] Seckler, M., Tuch, A., Opwis, K., and Bargas-Avila, J. User-friendly locations of error messages in web forms: Put them on the right side of the erroneous input field. *Interacting with Computers, Elsevier 24*, 3 (2012), 107–118.

[47] Shitkova, M., Holler, J., Heide, T., Clever, N., and Becker, J. Towards usability guidelines for mobile websites and applications. *Wirtschaftsinformatik* (2015), 1603–1617.

[48] Sorum, H., Andersen, K., and Vatrapu, R. Public websites and human–computer interaction: an empirical study of measurement of website quality and user satisfaction. *Behaviour and Information Technology 31*, 7 (2012), 697–706.

[49] Sweeney, S., and Crestani, F. Effective search results summary size and device screen size: Is there a relationship? *Information processing and management, Elsevier 42*, 4 (2006), 1056–1074.

[50] Wenger, J. On the rhetorical contract in humancomputer interaction. *Computers in Human Behavior, Elsevier 7*, 4 (1991), 245–262.

[51] Wroblewski, L. Web form design: filling in the blanks. *Rosenfeld Media* (2008).

[52] Wroblewski, L. Mobile first: Preface de jeffrey zeldmann. *Editions Eyrolles* (2012).

Appendix A

An Appendix

This section includes the questionnaires used in expirement one.

الاجابات على الأسئلة ادناه والمعلومات التي تم رصدها في تسجيل الفيديو يتم جمعها والإطلاع عليها لأغراض البحث العلمي فقط وبعد الانتهاء يتم اتلافها مباشرة. يتحمل الباحث مسؤولية اية استخدامات اخرى للمعلومات ويتعرض للمسائلة.

رقم المشارك:
الجنس: ( ) ذكر. ( ) أنثى.
العمر:
التخصص: ( ) تكنولوجيا معلومات. ( ) علم حاسوب . ( ) هندسة أنظمة حاسوب.
المستوى الجامعي: ( ) سنة اولى. ( ) سنة ثانية. ( ) سنة ثالثة. ( ) سنة رابعة. ( ) سنة خامسة.
اي الاجهزة تفضل استخدامها للدخول الى المواقع الالكترونية: ( ) الجوال. ( ) الحاسوب.
كم عدد الساعات التي تقضيها يومياعلى الانترنت:
( ) اقل من ساعة. ( ) من ساعة الى اثنتين. ( ) من ثلاث ساعات الى خمس. ( ) من ست ساعات الى ثمانية.

This section includes a questionnaire used in expirement Two.

الاجابات على الأسئلة ادناه والمعلومات التي تم رصدها في تسجيل الفيديو يتم جمعها والإطلاع عليها لأغراض البحث العلمي فقط وبعد الانتهاء يتم اتلافها مباشرة. يتحمل الباحث مسؤولية اية استخدامات اخرى للمعلومات ويتعرض للمسائلة.

القسم الأول: معلومات شخصية

رقم المشارك:

الجنس: ( ) ذكر. ( ) أنثى.

العمر:

التخصص: ( ) تكنولوجيا معلومات. ( ) علم حاسوب. ( ) هندسة أنظمة حاسوب.

المستوى الجامعي: ( ) سنة اولى. ( ) سنة ثانية. ( ) سنة ثالثة. ( ) سنة رابعة. ( ) سنة خامسة.

اي الاجهزة تفضل استخدامها للدخول الى المواقع الالكترونية: ( ) الجوال. ( ) الحاسوب.

مهارات استعمال الهاتف الذكي لتصفح الانترنت:

( ) سيئة. ( ) جيدة. ( ) جيدة جدا. ( ) ممتازة.

كم عدد الساعات التي تقضيها يومياعلى الانترنت:

( ) اقل من ساعة. ( ) من ساعة الى اثنتين. ( ) من ثلاث ساعات الى خمس. ( ) من ست ساعات الى ثمانية.

القسم الثاني: معلومات مرتبطة برسائل الخطأ في النموذج الذي يقوم المستخدم بتعبئته:

لم يكن مكان ظهور رسائل الخطأ مزعج:

( ) غير موافق بشدة. ( ) غير موافق. ( ) موافق. ( ) موافق بشدة.

لم يكن توقيت ظهور رسائل الخطأ مزعج:

( ) غير موافق بشدة. ( ) غير موافق. ( ) موافق. ( ) موافق بشدة.

كانت رسائل الخطأ مفيدة:

( ) غير موافق بشدة. ( ) غير موافق. ( ) موافق. ( ) موافق بشدة.

كانت الأخطاء سهلة التصحيح:

( ) غير موافق بشدة. ( ) غير موافق. ( ) موافق. ( ) موافق بشدة.

اعد مكان ظهور رسالة الخطأ افضل مكان:

( ) نعم. ( ) لا.

اذا كان الجواب على السؤال السابق ب لا، من تختار من التالية كأفضل مكان لظهور رسالة الخطأ:

( ) To the right of text box. ( ) To the left of text box. ( ) Above the label. ( ) Below label ( ) Below text box. ( ) popup message.

اعد وقت ظهور رسالة الخطأ افضل وقت:
( ) نعم. ( ) لا.
اذا كان الجواب على السؤال السابق ب لا، هل تفضل ظهور الرسالة مباشرة بعد الانتهاء من تعبئة الحقل:
( ) نعم. ( ) لا.

This section includes a questionnaire used in expirement Three.

الاجابات على الأسئلة ادناه والمعلومات التي تم رصدها في تسجيل الفيديو يتم جمعها والإطلاع عليها لأغراض البحث العلمي فقط وبعد الانتهاء يتم اتلافها مباشرة. يتحمل الباحث مسؤولية اية استخدامات اخرى للمعلومات ويتعرض للمسائلة.

القسم الأول: معلومات شخصية

رقم المشارك:

الجنس: ( ) ذكر. ( ) أنثى.

العمر:

التخصص: ( ) تكنولوجيا معلومات. ( ) علم حاسوب. ( ) هندسة أنظمة حاسوب.

المستوى الجامعي: ( ) سنة اولى. ( ) سنة ثانية. ( ) سنة ثالثة. ( ) سنة رابعة. ( ) سنة خامسة.

اي الاجهزة تفضل استخدامها للدخول الى المواقع الالكترونية: ( ) الجوال. ( ) الحاسوب. ( ) غير ذلك.

مهارات استعمال الهاتف الذكي لتصفح الانترنت:

( ) سيئة. ( ) جيدة. ( ) جيدة جدا. ( ) ممتازة.

كم عدد الساعات التي تقضيها يومياعلى الانترنت:

( ) اقل من ساعة. ( ) من ساعة الى اثنتين. ( ) من ثلاث ساعات الى خمس. ( ) من ست ساعات الى ثمانية.

القسم الثاني: معلومات مرتبطة برسائل الخطأ في النموذج الذي يقوم المستخدم بتعبئته:

اعد تظليل صندوق الادخال عند حدوث الخطأ امر مزعج:

( ) غير موافق بشدة. ( ) غير موافق. ( ) موافق. ( ) موافق بشدة.

اؤيد عملية تظليل صندوق الادخال عند حدوث الخطأ:

( ) غير موافق بشدة. ( ) غير موافق. ( ) موافق. ( ) موافق بشدة.

تظليل صندوق الادخال ادى الى ملاحظة مكان الخطأ بشكل اسرع:

كانت رسائل الخطأ مفيدة:

( ) غير موافق بشدة. ( ) غير موافق. ( ) موافق. ( ) موافق بشدة.

كانت الأخطاء سهلة التصحيح:

( ) غير موافق بشدة. ( ) غير موافق. ( ) موافق. ( ) موافق بشدة.