UNIVERSITY OF OKLAHOMA

GRADUATE COLLEGE

TOWARDS AN ENHANCED USER EXPERIENCE WITH SMART PHONE WEATHER ALERT APPLICATIONS: USABILITY AND USER-CENTERED DESIGN APPROACHES

A DISSERTATION

SUBMITTED TO THE GRADUATE FACULTY

in partial fulfillment of the requirements for the

Degree of

DOCTOR OF PHILOSOPHY

By

ABDULRAHMAN MOUSA KHAMAJ
Norman, Oklahoma
2018

TOWARDS AN ENHANCED USER EXPERIENCE WITH SMART PHONE WEATHER ALERT APPLICATIONS: USABILITY AND USER-CENTERED DESIGN APPROACHES

A DISSERTATION APPROVED FOR THE
SCHOOL OF INDUSTRIAL AND SYSTEMS ENGINEERING

BY

Dr. Ziho Kang, Chair

Dr. Hairong Song

Dr. Shivakumar Raman

Dr. Randa Shehab

Dr. Theodore Trafalis

Acknowledgment

My greatest thanks and appreciation must be given to all my family members for their infinite love and continuous support throughout the 6+ years of being abroad. My parents and siblings played a very important role in this achievement. My lovely wife, Alaa, and my hero son, Hussam, have beautified my life, greatly eased my studies, and immensely shaped who I am today.

This work would not have been possible without the guidance and support of my advisor, Dr. Ziho Kang. I highly thank Dr. Kang for everything he has done for me. I would also like to thank my committee members, Dr. Randa Shehab, Dr. Shivakumar Raman, Dr. Hairong Song, and Dr. Theodore Trafalis, for their valuable feedback and guidance throughout my doctoral program.

I am so grateful to all my friends for their encouragement, support, and good wishes. Special thanks for Mastoor, Hamoud, Yasser, Shabeeb, Abdulmajeed, Nawaf, Yahya, and Mustafa. In addition, I thank my lab mates, Amine, Saptarshi, and Salem, for their great help and advise throughout my entire research work.

Table of Contents

Acknowledgment iv

List of Figures xi

Abstract xvi

Chapter 1. Introduction 1

1.1 Overview 1

1.2 Problem Statement 5

1.3 Objectives 6

1.4 Research Questions 7

1.5 Chapter Summary and Dissertation Organization 8

Chapter 2. Literature Review 11

2.1 Smart Phone Technology 11

2.1.1 Current Trends and Future Prospects of Smart Phone App Market 11

2.1.2 Smart Phone Technology Among Different Age Groups 12

2.1.3 People Reliance on Smart Phone Weather Applications 15

2.2 Usability 16

2.2.1 Usability Definitions 16

2.2.2 Traditional Usability Evaluation Techniques 18

2.2.3 Eye Tracking-Based Usability Evaluation Technique 22

2.3 User-Centered Design 25

2.3.1 User-Centered Design Definitions 26

2.3.2 User-Centered Design Processes and Evaluation Techniques 27

2.4 Previous Usability and User-Centered Design Research of Smart Phone Weather Applications 30

2.5 Chapter Summary 34

Chapter 3. A Mixed Methods Approach to Evaluating the Usability of Smart Phone Weather Alert Applications 36

3.1 Introduction.. 36

3.2 Study I: Traditional Usability Analysis of Smart Phone Weather Applications... 43

3.2.1 Objective 43

3.2.2 Method... 43

3.2.3 Results 52

3.2.4 Discussion of Study I 62

3.3 Study II: Eye Tracking Analysis of Smart Phone Weather Applications............. 66

3.3.1 Objective 66

3.3.2 Method... 66

3.3.3 Results 71

3.3.4 Discussion of Study II 96

3.4 Chapter Summary 101

Chapter 4. User-Centered Design Assessment of End-User Needs in Smart Phone Weather Applications (Phase 1) 104

4.1 Introduction.. 104

4.2 Users' Perceptions of Smart Phone Weather Applications' Usability: A Descriptive Qualitative Assessment 105
4.2.1 Method 106
4.2.2 Results 109
4.2.3 Discussion 116
4.3 Usability Heuristics for Smart Phone Interface Designs 117
4.3.1 General Essential Usability Heuristics 118
4.3.2 Specific Usability Heuristics for Smart Phone Older Users 122
4.3.3 Smart Phone Application Design Heuristics 124
4.4 Chapter Summary 127
Chapter 5. Development of A Smart Phone Prototype Weather Application Based on the User-Centered Design Approach (Phase 2) 128
5.1 Introduction 128
5.2 Structure and Content of the Prototype Application 128
5.3 Chapter Summary 146
Chapter 6. Usability Evaluation of Older and Younger Users' Experiences with the Smart Phone Prototype Weather Application (Phase 3) 148
6.1 Introduction 148
6.2 Method 149
6.3 Results 154
6.4 Discussion 173

6.5 Chapter Summary 177

Chapter 7. Summary, Recommendation, and Conclusion........................ 179

7.1 Research Summary 179

7.2 Research Contribution.. 184

7.3 Practical Recommendation 186

7.4 Limitations and Future Research 190

7.5 Conclusion 191

References.. 193

Appendix A: Exit Survey Questionnaire Used in Chapter 3 208

Appendix B (1): Original and proposed STW messages Used in Chapter3 209

Appendix B (2): Original and proposed WA messages Used in Chapter3 210

Appendix C: Focus Group Questions Used in Chapter 4 211

List of Tables

Table 1. Survey statements for the original and proposed STW messages 49

Table 2. Survey statements for the original and proposed WA messages 50

Table 3. Mixed design ANOVA output for the STW weather alert messages 57

Table 4. Mixed design ANOVA output for the WA weather alert messages 59

Table 5. Content analysis of usability comments from both first-time and experienced users 61

Table 6. Steps for performing a thematic analysis, summarized from Braun and Clarke (2006) 109

Table 7. Focus group findings (themes, sub-themes, and frequency of occurrence) ... 110

Table 8. Nielsen’s severity rating scale of the usability problems 152

Table 9. Mann-Whitney test summary for number of errors 157

Table 10. Causes of errors, frequency of issue, (proportions of users who made errors), and average severity ratings on Weather Radio 158

Table 11. Causes of errors, frequency of issues, (proportions of users who made errors), and average severity ratings on EZ Weather 159

Table 12. Two-way ANOVA summary for task completion time 161

Table 13. Mann-Whitney test summary for post-task satisfaction ratings 164

Table 14. QUIS overall reaction questions 165

Table 15. QUIS screen questions 166

Table 16. QUIS terminology & system information questions 167

Table 17. QUIS learning questions 169

Table 18. QUIS system capabilities & multimedia questions 170

Table 19. Correlation matrix for younger and on Weather Radio (*r* & (*p*-values)) 172

Table 20. Correlation matrix for older and on Weather Radio (*r* & (*p*-values)) 172

Table 21. Correlation matrix for younger and on EZ Weather (*r* & (*p*-values)) 172

Table 22. Correlation matrix for older and on EZ Weather (*r* & (*p*-values)) 173

List of Figures

Figure 1. (a) Simple example of scanpath structured by 6 eye fixations (size of eye fixations is proportional to the duration spent on that fixation) and saccades (transitions between fixations), mapped into 3 AOIs. (b) A transition matrix derived from (a) 23

Figure 2. Framework of user-centered design 29

Figure 3. Example of location search process on Weather Radio 37

Figure 4. Example of accessing weather forecasts process on Weather Radio 39

Figure 5. Example of accessing alert messages process on Weather Radio 40

Figure 6. Example of controlling alert settings process on Weather Radio 41

Figure 7. Example of controlling radar map settings process on Weather Radio 42

Figure 8. Plot of task completion time for the location search; POM: pin on map and ATB: app's text bar 53

Figure 9. Plot of task completion time for the alert settings 53

Figure 10. Plots of task completion time for the map settings 54

Figure 11. First-time and experienced users' mean rating for each STW survey item: (a) header information, (b) use of delimiter, (c) letter format, and (d) overall satisfaction . 56

Figure 12. First-time and experienced users' mean rating for each WA survey item: (a) information location, (b) word expression, (c) use of terminology, and (d) overall satisfaction 58

Figure 13. Current method of delivering alert messages to end users 65

Figure 14. Experimental setup showing an example of how a user's eye fixation (yellow circle) on the actual smartphone appear on the emulated interface with the same distance (d) from a reference point 68

Figure 15. Example of overall and specific AOIs of the Weather Radio app 71

Figure 16. Differences between first-time and experienced users on the location search task in terms of the overall eye fixations duration (a) and number (b)........ 71

Figure 17. Differences between first-time and experienced users on the alert settings task in terms of the overall eye fixations duration (a) and number (b)........ 73

Figure 18. Differences between first-time and experienced users on the map settings task in terms of the overall eye fixations duration (a) and number (b)........ 74

Figure 19. Differences between first-time and experienced users on location search approaches: POM (a & c) and ATB (b & d), in terms of the mean eye fixations duration (a & b) and number (c & d) 76

Figure 20. Differences between first-time and experienced users on the alert settings task, in terms of the mean eye fixations duration (a & b) and number (c & d) 78

Figure 21. Differences between first-time and experienced users on the map settings task, in terms of the mean eye fixations duration (a & b) and number (c & d) 80

Figure 22. Example of a user's accumulative scanpath on the alert settings task 82

Figure 23. Scanpath pattern needed to successfully complete the location search (pin on map) 83

Figure 24. First-time (a) and experienced (b) users' eye transitions among AOIs for location search (pin on map)........ 84

Figure 25. Scanpath pattern needed to successfully complete the location search (app's text bar) ... 85
Figure 26. First-time (a) and experienced (b) users' eye transitions among AOIs for location search (app's text bar) ... 85
Figure 27. Scanpath pattern to successfully complete the alert settings task ... 86
Figure 28. First-time (a) and experienced (b) users' eye transitions among AOIs for alert settings task ... 87
Figure 29. Scanpath pattern to successfully complete the map settings task ... 87
Figure 30. First-time (a) and experienced (b) users' eye transitions among AOIs for map settings task ... 88
Figure 31. Scatterplots of the correlations between (a) fixation durations and numbers, (b) fixation durations and transition numbers, and (c) fixation numbers and transition numbers for first-time users on the location search task (pin on map) ... 90
Figure 32. Scatterplots of the correlations between (a) fixation durations and numbers, (b) fixation durations and transition numbers, and (c) fixation numbers and transition numbers for experienced users on the location search task (pin on map) ... 91
Figure 33. Scatterplots of the correlations between (a) fixation durations and numbers, (b) fixation durations and transition numbers, and (c) fixation numbers and transition numbers for first-time users on the location search task (app's text bar) ... 92
Figure 34. Scatterplots of the correlations between (a) fixation durations and numbers, (b) fixation durations and transition numbers, and (c) fixation numbers and transition numbers for experienced users on the location search task (app's text bar) ... 93

Figure 35. Scatterplots of the correlations between (a) fixation durations and numbers, (b) fixation durations and transition numbers, and (c) fixation numbers and transition numbers for first-time users on the alert settings task 94
Figure 36. scatterplot of the correlation between fixation duration and fixation numbers for experienced users on the alert settings task 94
Figure 37. Scatterplots of the correlations between (a) fixation durations and numbers, (b) fixation durations and transition numbers, and (c) fixation numbers and transition numbers for first-time users on the map settings task 95
Figure 38. scatterplot of the correlation between fixation duration and fixation numbers for experienced users on the map settings task... 96
Figure 39. Example of clear call to action on installation screen................................ 129
Figure 40. Example of optimized features with the use of appropriate visualization and swiping functionality 130
Figure 41. Example of appropriate use of labeled icons and descriptive information . 132
Figure 42. Example of consistent and minimalist design with a “back” feature.......... 133
Figure 43. Example of flexible and efficient feature, easy manual location change, and effective search index 134
Figure 44. Example of a visible system status and user control & freedom 135
Figure 45. Example of visible and intuitive feature 136
Figure 46. Example of easy language, structured information, clear color contrasts, and large text font sizes 137
Figure 47. Example of location search process on EZ Weather.................................. 139

Figure 48. Example of accessing weather process on EZ Weather 140

Figure 49. Example of accessing alert messages process on EZ Weather 142

Figure 50. Example of controlling alert settings on EZ Weather 144

Figure 51. Example of accessing radar map and controlling map settings on EZ Weather 145

Figure 52. Proportions of successful task completion for both age groups on (a) Weather Radio and (b) EZ Weather 155

Figure 53. Mean number of errors for both age groups on all tasks (a to e) 156

Figure 54. Mean task completion time spent in completion of all tasks (a to e) 160

Figure 55. Mean post-task satisfaction ratings of SEQ survey for all tasks (a to e)..... 163

Figure 56. Users' mean satisfaction ratings for the overall reaction category on Weather Radio and EZ Weather.. 165

Figure 57. Users' mean satisfaction ratings for the screen category on Weather Radio and EZ Weather 166

Figure 58. Users' mean satisfaction ratings for the terminology & system information category on Weather Radio and EZ Weather 168

Figure 59. Users' mean satisfaction ratings for the learning category on Weather Radio and EZ Weather 169

Figure 60. Users' mean satisfaction ratings for the system capabilities & multimedia category on Weather Radio and EZ Weather 171

Abstract

Today, smart phones are ubiquitous in our everyday lives. We rely heavily on their immediately available features, especially when time-critical and/or life-saving information, such as weather alert messages, need to be easily and quickly accessed. With the advancement of smart phone technology, smart phone weather alert applications (apps) have been continuously developed and launched to the market. However, many app developers may pay more attention to creating various features and highly sophisticated tools than considering the most important factors of the design- the usability and users' needs. Overlooking usability principles and end-users' needs in the design phase of any system interface can be associated with effectiveness, efficiency, and user satisfaction issues. More importantly, poor usability of time-critical interfaces (e.g. smart phone weather app interfaces) may even hinder performing life-saving actions. So far, only little attention has been devoted to the usability evaluation and end-user needs with weather alert apps on the smart phone platform.

To address this issue, the work in this dissertation is centered on performing systematic usability and user-centered design (UCD) analysis approaches to evaluate and enhance the usability of smart phone weather apps, with a specific focus on first-time users. Specifically, (1) using both traditional (i.e. task completion rate, task completion time and performance surveys) and eye tracking (i.e. eye fixation durations, eye fixation numbers, and scanpath observations) measures, we evaluate the usability of smart phone weather apps with the goal of identifying usability problems; (2) using focus group interviews, we investigate end-users' goals and needs in weather apps (first

UCD phase); (3) considering the focus group findings, general usability heuristics, specific user groups' (i.e. older users) limitations and recommendations, and smart phone app design principles, we develop a prototype smart phone weather app (second UCD phase); and finally (4) using both quantitative and qualitative evaluation approaches, we evaluate the developed prototype app to validate its usability (last UCD phase). Findings revealed multiple usability problems with currently running smart phone weather apps and showed that the developed prototype app that followed the UCD approach, greatly enhanced users' experiences of different age groups compared to a representative popular weather app. We make several recommendations for future designs of smart phone weather apps, as well as apps that share similar features and characteristics.

Chapter 1. Introduction

1.1 Overview

Being weather-aware is extremely important to one's safety, especially during seasons known to have hostile weather conditions, such as floods, tornados, hurricanes, and heat exhaustion in certain geographical areas. This importance is supported by the frightening statistics about the consequences of weather-related incidents. According to the most recent statistics from the National Weather Service (NWS), weather-related events in the United States only, caused 508 fatalities, 1205 injuries, and around $90 million in damage costs in 2017 (National Weather Service, 2018). Hence, communication of daily weather forecasts and weather alert notifications issued by authorized sources such as the National Oceanic and Atmospheric Administration (NOAA) and the NWS and operated by private agencies should efficiently help people comprehend the weather situation and be prepared/react accordingly.

There are several sources that deliver weather information to the public, such as through television, radios, and smart phones. However, with the advancement of technology and the increasing use of smart phone devices, a great number of people rely heavily on smart phone weather alert apps in accessing weather information. Zabini (2016) stated that people utilize smart phone apps for weather information more than all other information sources. In 2018, statistics revealed that smart phone weather apps were among the seven most used smart phone app categories in the United States, with more than 91% smart phone reach among users (Statista, 2018a). The people's tendency in use of smart phone weather apps is possibly because they are readily available at users' fingertips and can easily and quickly enable them to access weather information in different contexts of use (while on the move, eating, talking...etc.).

Even though NOAA and NWS are striving to produce reliable weather forecast information in a timely manner, the capability of users to easily interact with such information on smart phone weather apps is of great focus in this dissertation. This is mainly due to a general tendency of numerous smart phone app developers, in the design stage, to focus more on creating as many features as possible for their users rather than to ease users' interaction and exactly meet their needs. In order to satisfy intended users and grow business, the usability of such apps must be evaluated with the aim of identifying and solving usability problems (Hussain et al., 2017; Hussain et al., 2015; Harrison et al., 2013; Hussain et al., 2017; Williams, 2004; Hussain & Mkpojiogu, 2015).

Usability is defined as the capability of the user to understand, learn, and use an interface as well as perceive it as an attractive under different conditions (International Organization for Standardization (ISO 9126-1), 2001). Ignoring the usability of interfaces, especially those containing time-critical data (e.g. weather app interfaces), and not thoroughly considering the different characteristics of the ultimate users, may lead to severe consequences. Poorly designed apps may fail to convey the weather alerts (the risk level associated with the weather feature) properly; especially during severe weather situations that require appropriate reaction in a timely manner.

One typical scenario of using smart phone weather apps is that many users, upon installing their apps, may not take the time to fully learn how the interfaces are manipulated, but rather they may keep the default settings (e.g. inactivated alert notifications and auto-location detection). There can be extreme cases of weather conditions at a particular location and a user needs to immediately access relevant information in order to make safety decisions. For example, if a user living in the state of Maine becomes aware through a radio that a huge tornado is forming in the state of Oklahoma, where their grandmother lives, they may need to urgently: 1) add their grandmother's location; 2) access relevant weather forecasts (e.g. humidity, wind speed, and chance of rain); 3) change the default map settings (e.g. change map type from "standard" to "satellite" view to verify if her location is in the tornado zone); 4) activate the tornado alerts (e.g. tornado warning and watch alerts); 5) receive and read alert notifications/messages pushed by local agencies about her location. If such features are difficult to use, users may not be able to become fully aware of the risk level associated with the time-critical weather condition, resulting in serious consequences.

As the user is the target and main part in any interface design process, researchers have continuously called for considering a user-centered design (UCD) approach when developing or designing system and product interfaces (Garrett, 2010; Lack, 2007; Vredenburg et al., 2002; Abras et al., 2004). The UCD refers to iterative steps that are centered around the needs and limitations, and characteristics of end-users by carefully considering them in each step (Mao et al., 2005). When an interface developer or designer well understand their intended users, and then design from users' perspective, the interface has a great chance to be usable, leading to successful business (Gladkiy, 2018).

Apart from user needs in any interface design, limitations and characteristics of specific end-user groups must also be a top priority for developers. One important example is the older users' group. Age-related changes typically arise when people reach the age of 50 years (Wahrendorf et al., 2013). Older users (50+ years old) generally suffer from problems that arise with age, such as decrease in working memory and visual, cognitive & motor capabilities (Wahl & Römer, 2001; Lawton, 1990; Sekuler & Ball, 1986; Sweller, 1988; Czaja et al., 2006). In addition to their physical and cognitive difficulties, older users have difficulty in coping with the rapid progression of smartphone technology (Khawaji, 2017). Any usability issues with poorly designed interfaces and their consequences are expected to be worse for this important user group, compared to other user groups such as younger users. The consequences would even be worse if the interface includes time-critical data such as that of weather app interfaces. Hence, it is very crucial to design an interface with users' limitations and different characteristics in mind.

To uncover as many usability issues as possible, it is often recommended to test the design by novices. Throughout the past years, researchers (Bourie et al., 1997; Donker and Reitsma, 2004; Gerardo, 2007; Faulkner and Wick, 2005; Kjeldskov et al., 2005) concluded, after performing multiple deep analyses, that novices encounter much more usability problems than experts. System interfaces should be easy to use from the first time of use by any user; if not, then novices are likely to suffer more than experts, regardless of experts' experience levels (Faulkner and Wick, 2005). Hence, in this dissertation, we mainly focus on first-time users for discovering usability issues and needs, as well as building new enhanced interfaces.

1.2 Problem Statement

Usability research has been well established in the human factors field with a focus on several different domains such as in health (Segall et al., 2011), aviation (Clamann and Kaber, 2004), in-vehicle infotainment systems (Khamaj et al., 2017), and virtual environment (Bowman, 2002). In addition, several studies have focused on examining the usability of smart phone apps in different specific areas such as in tourism (Geven et al., 2006; Shrestha 2007; Ahmadi and Kong 2008; Schmiedl et al., 2009) and geography (Elzakker et al., 2008). What is not well grounded in current research, however, is evaluating the usability of smart phone apps with embedded time-critical and/or life-saving information. In particular, smart phone weather apps have been considered for usability evaluation in only few studies with even shallow investigation.

Even though Singhal (2011), Alluri (2012), and Drogalis et al., (2015) examined the usability of weather apps based on a few tasks given to participants, their findings may not be reliable nor generalizable to other experimental settings. Specifically, Singhal (2011) and Alluri (2012) tested the usability of all native apps, including weather apps, in iPhone and Android, respectively. Apart from the fact that they evaluated user performance on only one or two tasks as well as that weather apps were not the main focus for evaluation, native weather apps contain only basic weather information and may not be the essential sources for most users. Furthermore, the pilot study by Drogalis et al. (2015) included only six participants and did not consider any benchmark approach or credible standard criteria to be used as reference measures for their collected data.

In addition to the aforementioned issues, current smart phone weather app usability research only employs conventional evaluation metrics (e.g. task completion time and satisfaction surveys). Though these metrics can elicit great knowledge about the usability of interfaces, they may not well inform us about users' cognitive processes and decision-making strategies during their interaction with interfaces. Even knowledge elicitation tools (e.g. think aloud protocols and interviews) that are aimed to solve this issue, may not be completely reliable because of issues with tacit users' answers (Chervinskaya & Wasserman, 2000). A great way to answer how and why users interact with an interface the way they do is by utilizing the physiological measurement tool, eye tracking (Jacob and Karn, 2003). However, existing usability research of smart phone weather apps indicates a gap in this area.

Another gap exists in the knowledge related to this area is the use of the UCD approach throughout the development stages of real-time smart phone apps such as weather apps. Findings from several domains suggest that the UCD technique may enhance end-user performance and satisfaction and contributing to business growth (Morey et al., 2017).

1.3 Objectives

The main objective of the work in this dissertation is to add breadth to the literature and account for existing research limitations regarding the usability of smart phone weather alert apps with the focus on enhancing experiences of first-time users. First, we perform preliminary evaluation of smart phone weather apps' usability by taking one of the most popular and widely used weather apps, Weather Radio, as a case study. Specifically, we perform the evaluation using both conventional usability evaluation metrics: task time and surveys to get an overall picture of how users interact with weather apps for the first time. In addition, we use a modern evaluation tool: eye tracking to objectively understand users' cognitive processes and decision-making strategies when performing given tasks. The goal of the preliminary evaluation is to objectively know the usability issues that arise from first-time users' interaction with weather apps and to provide a base for further investigation.

Second, we employ the UCD approach to determine if it will improve users' performance on and perception of weather apps. Specifically, we: (1) perform focus group interviews to discuss regular users' goals, limitations, and needs in smart phone weather apps as well as explore the most popular and crucial usability and smart phone app design heuristics for users of different age groups; (2) we develop a prototype smart phone weather app that is fully designed based on users' inputs and widely accepted heuristics from (1); (3) we evaluate the usability of the newly built prototype app among different age groups: younger and older users, and by benchmarking it with a widely used weather app.

1.4 Research Questions

This dissertation aims to answer three interrelated research questions with the joint goal of adding knowledge to usability evaluation with first-time users, user-centered design, smart phone apps of real-time data, and smart phone apps of weather data. Four studies, detailed in the next chapters, used objective and subjective research approaches to answer these questions. The research questions are as follows.

- What are the usability problems in current smart phone weather apps?
- How does eye tracking support conventional evaluation metrics in informing the usability of smart phone weather apps?
- To what extent will employing the UCD approach, as well as considering key usability principles in upcoming weather app designs, improve user performance and satisfaction on all features, regardless of age?

1.5 Chapter Summary and Dissertation Organization

This chapter introduces the current research and illustrates the rationale for examining the usability of existing smart phone weather apps as well as the need for building a new app from the user point of view. Specifically, the chapter emphasizes that usability evaluation is a crucial step for the success of any system or product, especially those containing time-critical information, such as weather information. It also shows that the UCD approach of learning about user goals, needs, characteristics, and limitations; then designing an interface from their perspective will result in both enhanced user experience and business success.

This dissertation consists of seven chapters. Each chapter's content is briefly described as follows. Following Chapter 1, Chapter 2 provides background information about smart phone technology and its use among different age groups with a particular focus on the weather domain. It also discusses usability and UCD theories, processes, evaluation methods, and their previous research. In Chapter 3, using a mixed methods approach of traditional and highly analytical objective metrics, we test the usability of smart phone weather apps among first-time users. Chapter 4 contains the first phase of the UCD process, which is a qualitative focus group assessment of user feedback about the usability of smart phone weather apps and heuristic guidelines for usable interfaces. The analyses in this chapter greatly help understanding end-users' requirements and how they can be implemented in future designs according to usability heuristics and smart phone app design specifications. Following this chapter, Chapter 5 includes the second phase of the UCD process, which presents the structure and content of a prototype smart phone weather app designed based on the findings from Chapter 4. Chapter 6 includes the third and last phase of the UCD process, which evaluates the usability of the UCD-based prototype weather app. The evaluation includes both younger and older users to determine if the prototype is user friendly for all users, regardless of age.

Most of the content in this dissertation is either published in scientific journals or submitted/to be submitted soon to prestigious journal or conference societies. Specifically, Section 3.2 in Chapter 3 is from the published paper titled "Usability evaluation of mobile weather hazard alert applications" in the journal of *Industrial and Systems Engineering Review* (Khamaj and Kang, 2018). Section 3.3 in Chapter 3 is from the submitted paper titled "Integrated eye movement analysis approach of time-critical services: application in a mobile weather alert system" to the journal of *Industrial Engineering and Management Systems*. Section 4.2 in Chapter 4 is from the draft of the paper titled "Users' perceptions of mobile weather applications' usability" that will be submitted soon to the 2019 annual *Human Factors and Ergonomic Society (HFES)* conference. Chapter 6 is from the draft of the paper titled "Usability evaluation of time-critical weather alert application through features characterized from user-centered design process" that will be submitted soon to the journal of *Human Factors and Ergonomic Society (HFES)*. The contents of these papers are elaborated and explained in more details in this dissertation.

Chapter 2. Literature Review

In this chapter, we discuss the smart phones usage among people in the current era and the trend towards their use in the future. Also, we focus on peoples' increasing reliance on accessing weather information through smart phones. Then, we explain the importance of investigating usability, its definitions, and its widely used evaluation methods. Following this, we illustrate the need for the UCD approach, its definitions, and its processes & evaluation techniques. Finally, we discuss previous research regarding the usability and UCD of smart phone weather apps and existing gaps.

2.1 Smart Phone Technology

Unlike bulky computer devices such as desktop computers, a smart phone is defined as "a hand-held computer capable of multiple functions in addition to placing calls" (Kaplan, 2012, p.129). In addition, smart phones are typically touchscreen-based interfaces that enable users to get access to the internet, browse websites, and download and operate software apps. As the definition informs, smartphones tend to make lives easier as users can get full and easy access to the technology through their devices in multiple different contexts of use, such as while eating, walking, etc. (Nayebi et al., 2012). Smart phone users are continuously increasing worldwide. For example, recent statistics showed that more than 2 billion people globally used smartphones in 2017, with around 224 million users during the same time in the United States alone; the number is estimated to increase rapidly (Statista, 2018c).

2.1.1 Current Trends and Future Prospects of Smart Phone App Market

The smart phone app market is growing so fast; the industry is getting bigger on a daily basis; developers' population has been astonishingly increasing (Stacy, 2017). The increasing number of smartphone users has encouraged companies and technology experts to develop vast numbers of apps to be used by smart phone consumers. For instance, more than 150,000 apps were available for Android users only and around 350,000 users activate apps daily (Xu et al., 2011). In 2017, the number of smart phone app downloads worldwide was about 178 billion and was forecasted to reach 205.4 billion and 258.2 billion by the end of 2018 and 2022, respectively (Statista, 2018c).

Researchers are in line with the revolution in smart phone app technology. For example, Hussain and Kutar (2009) believe that all peoples' life matters, especially those of business, should be accessed through smart phone apps in order to cope with the increasing demand of these apps as well as to maintain a position in market. The smart phone app market has been utilized in several domains, such as in health and medical emergency (Chittaro et al., 2007; Holzinger and Errath, 2007), disasters (Lee et al., 2011; Monares et al., 2011), and transportation (Argyle et al., 2015). In addition, Hussain and Kutar (2009) claim that smart phone apps containing real-time information such as urgent news and weather apps have been very common and highly used.

2.1.2 Smart Phone Technology Among Different Age Groups

With the proliferation of smart phone app's market, people of different age groups have been increasingly using smart phone apps. Two important age groups are of particular interest in this dissertation: younger and older adults. Statistics showed that the percentage of younger adults (aged 18 - 29 years old) in the United States who owned smart phones increased from 86% in 2015 to 94% in 2017 (Statista, 2018b). A larger smart phone ownership percentage increase was observed among older adults (50+ years old). For those aged 50 – 64 years old, the smart phone ownership percentage increased from 58% to 73%, and from 30% to 46% for those aged 65+ years old.

This huge increase, especially among older adults, clearly implies that all people of different age groups are either willing to keep up with the advancement in technology or are forced to adopt it. Either one, it is extremely important that the interaction between users and smart phones is in the best way possible. To achieve that, users' characteristics, needs, and limitations must be taken into consideration when designing interfaces such as downloadable smart phone apps.

While smart phone usage among younger adults is easy and perceived as useful since they grow up together (Pan et al., 2013), natural barriers that arise as people become older may make older adults feel less independent and live in a low quality of life (Barros et al., 2014; Dix et al., 2004). In addition, previous studies found that smart phone interfaces are mainly developed for younger adults (Fisk et al., 2009; Lorenz & Oppermann, 2008). The barriers among older adults are often categorized under two important factors. These factors are explained as follows.

Cognitive Factors

Previous research illustrated that older users tend to suffer from declining in their cognitive abilities and working memory, which in turn make it very challenging to acquire and recall new skills and to perform several tasks (Czaja et al., 2006; Fisk et al., 2009; Leung et al., 2010). With the rapid progression of technology and continuous tendency of developers to build all-inclusive interfaces that contain large amount of information and features, cognitive load issues might arise with larger negative impact on older users. Cognitive load refers to how much information can be stored in a person's working memory at once (Sweller, 1988). When cognitive load increases during the interaction with an interface, all users, especially older users, are likely to not easily find a relative information, experience delayed decisions, and make errors (Adcock, 2000).

Physical and Perceptual Factors

Physical and perceptual limitations that older users may experience include declining in motor functions, decreases in visual and auditory capabilities, and reduced mobility (Charness et al., 2001). However, vision is the most important physical and perceptual factor in the interaction between users and system interfaces. As visual abilities decrease, it is expected for older users to have difficulty interacting with interfaces containing small font sizes, too much clutter, and low color contrasts (Bitterman & Shalev, 2004; Fisk et al., 2009; Kurniawan, 2008; Pak & McLaughlin, 2010).

In addition to these factors, Leung et al. (2010) highlighted an important barrier for older users, compared to younger users, which negatively impact their confidence and independence: little smart phone technology experience. Moreover, Fisk et al. (2009) found that older users struggle a lot more than younger users in adopting new technologies and interacting with enhanced features due to their little previous knowledge and skills; this is also reflected on their mental models, where they may perceive technology differently from younger users.

2.1.3 People Reliance on Smart Phone Weather Applications

Delivering weather information to the public is considered one of the most crucial tools for safety and awareness, with respect to natural calamity. Daily weather forecasts and weather alert notifications sent by authorized sources such as the NWS play an important role in alerting people about potential hazards and making decisions about outdoor activities. The means for conveying these predictions should be very efficient and accessible.

As explained in Chapter 1, there are multiple sources that deliver weather information to the public such as televisions, radios, and smartphones. Zabini (2016) found that with the revolution of technology, people tend to utilize smart phone apps for weather information more than all other information sources, including those mentioned above. Statistics are in line with Zabini's (2016) findings, as it was revealed in 2018 that smart phone weather apps are among the seven most used smart phone app categories in the United States, with more than 91% smart phone reach among users (Statista, 2018a). In addition, more than 8000 weather apps are available in the iTunes app store alone, as of August 2018 (iTunes, 2018) and over 5.2 million users have installed and run the Weather Radio app, created by Weather Decision Technologies (WDT), on their devices (Weather Decision Technology, 2018).

Due to the accessibility of advanced app software development tools, any developer can easily create new weather apps. Statistics showed that the number of active weather apps in the iTunes App Store alone increased from 5,043 apps in December 2014 to 8,006 apps as of June 2018 (iTunes, 2018). This rapid increase in weather apps may come at the expense of the usability of these apps. In other words, these apps significantly vary in quality as many app developers, in the design stage, tend to pay less attention to the importance of the "ease of use" factor, and rather focus on creating as many features as possible. If weather app interfaces are not well designed, users may not effectively, efficiently find and process the presented information, resulting in poor user experience, business loss, and perhaps leading to unfortunate consequences. Hence, it is very important for users to be able to easily interact with weather app features, especially during natural disasters that require prompt and appropriate reactions.

2.2 Usability

To examine the nature of interaction between users and interfaces (weather interfaces in this work), recommend improvement modifications, and/or build new enhanced interfaces, we need to delve into the so-called field: "usability".

2.2.1 Usability Definitions

Several definitions of usability are available in the literature; perhaps due to its long-established in the Human Factors (HF) and Human Computer Interaction (HCI) fields of research. Shackel (1991) and Preece et al. (1994) have introduced a comprehensive view of usability. Specifically, Shackel (1991) defined usability as "the capability in human functional terms to be used easily and effectively by the specified range of users, given specified training and user support, to fulfill the specified range of tasks, within the specified range of environmental scenarios" (p. 24). In addition, Preece et al. (1994) thought of usability as "a measure of the ease with which a system can be learned or used, its safety, effectiveness and efficiency, and attitude of its users towards it" (p.722). The Institute of Electrical and Electronics Engineers (1993) also defined usability as "the ease with which a user can learn to operate, prepare inputs for, and interpret outputs of a system or component." Similarly, the International Organization for Standardization (ISO 9126-1) (2001) later defined usability as "the capability of the software product to be understood, learned, used, and attractive to the user, when used under specified conditions."

To determine the usability, a usability evaluation technique should be employed. Usability evaluation is a very critical step to the success of any product or system (Nielsen, 1994a). Scholars in the smart phone platform research also pointed out that usability evaluation is one of the most important techniques to test the quality and discover the challenges and limitations within smart phone apps (Baharuddin et al., 2013). Usability evaluation can be defined as a set of procedures used for evaluating the usability and identifying issues that result from the interaction between users and a system's interface design (Saleh & Ismail, 2015). Similarly, other scholars define usability evaluation as processes used to determine the product's current level of usability, identify usability problems, and suggest reliable solutions to the identified problems (Nielsen, 1994a; Lettner & Holzmann, 2011; Ball & Bothma, 2017). Findings from usability evaluation can significantly help developers improve their systems/products, which ultimately lead to users' satisfaction and success in businesses (Lizano et al., 2013; Alshehri & Freeman, 2012).

2.2.2 Traditional Usability Evaluation Techniques

There are several usability evaluation methods considered in the literature. Each method can be measured with two kinds of data: quantitative and qualitative, depending on the goal of the evaluator and the context of use; sometimes both kinds of data are employed. Direct observational method (Khanum & Trivedi, 2012; Tangsoc & Amelia, 2009), focus group method (Krueger & Casey, 2002; Goodman et al., 2004), GOMS method (John & Kieras, 1996), heuristic evaluation (Gómez et al., 2014; Inostroza et al., 2013), and think aloud protocol (Jones et al., 2017) are among the most popular and frequently used traditional usability evaluation methods.

Two of the aforementioned evaluation methods are advocated by the most cited sources in the usability and HCI fields: direct observational method (International Organization for Standardization (ISO 9241-11), 1997; Nielsen, 1994a) and focus groups (Nielsen, 1997). These two methods are extensively used in this dissertation to well inform the usability of smart phone weather apps. The direct observational method's definition and attributes with supporting citations from the literature are detailed as follows. Next, the focus groups method is explained.

Direct observational method refers to any procedure used to observe the performance of a user on a given task (Barendregt et al., 2003). This method includes six frequently used attributes for measuring the usability of an interface. These attributes and what they measure are outlined as follows.

Effectiveness

Effectiveness can be defined as the extent to which a user is able to perform a task in a specified context (International Organization for Standardization (ISO 9241-11), 1997; Shackel, 1991; Nielsen,1994a). Effectiveness has been well known to be measured by whether or not a user can successfully complete a given task and quantified by a task successful completion rate or percentage of users who fail to accomplish a task (Harrison et al., 2013). On the other hand, Frøkjær et al. (2000) believe that effectiveness is best to be measured through the user's outcome quality of interaction with an interface and quantified by the number of requests for assistance to successfully complete a given task. Hence, an interface is said to be effective if it attains high rates of successfully completed tasks and a small number of assistance requests.

Efficiency

Efficiency refers to what extent a user accurately completes a certain task with the assistance of expended resources in order to achieve a designated goal (Frøkjær et al., 2000). This attribute is widely used in the literature as it reflects the level of a user's accuracy when interacting with an interface. For example, Nielsen (1994a), Constantine and Lockwood (1999), and Seffah et al. (2006) stated that "efficiency in use" is among the most useful attributes, which describes the quality of an interface. Similarly, Schneidrman (1992) referred to efficiency as how quickly a user performs a certain task. In addition, Preece et al. (1994) considered efficiency as one of the most informative attributes about the productivity of a user on a given task. The usability in regard to this attribute can be measured in various ways depending on the context of use; however, most usability researchers consider the task completion time as the most appropriate indicator of efficiency. An efficient interface is the one that requires less task completion time.

Learnability

Harison et al. (2013) defines learnability as "the ease with which a user can gain proficiency with an interface" (p.4). Specifically, how fast a new user can engage with a system with efficiency and flawless interaction is the main focus of learnability. Schneidrman (1992) and Shackel (1991) quantified the learnability attribute as the time required for a user to learn how to use a system. Even though there are various ways to gauge the learnability attribute, such as number features learned, re-leaning time, and user's subjective opinion regarding ease of learning, the time to learn using a new interface is the most common measure for learnability. If the learning time is short and/or perceived as short, the system is considered to be an easy to learn.

Memorability

Memorability can be defined as how easy a user can successfully accomplish a system's task after a period of not using the system (Nielsen, 2012). To quantify memorability, Nielsen (2012) suggests using the number of trials and/or the amount of time needed to successfully accomplish a task through a repeated exposure to an interface. A memorable interface is the one that requires a smaller number of trials and less amount of time on a time after a time of being away.

Errors

This attribute can be described in terms of the number of errors made by users, severity of the errors, and the ability to recover from the errors during the interaction with a system (Nielsen, 2012). As an indicator of simplicity and usability, Schneidrman (1992) believes that calculating the number of errors made by a user or their error rate during the interaction with an interface would well serve this purpose. And of course, the smaller number of errors on an interface, the simpler and usable the system is.

User satisfaction

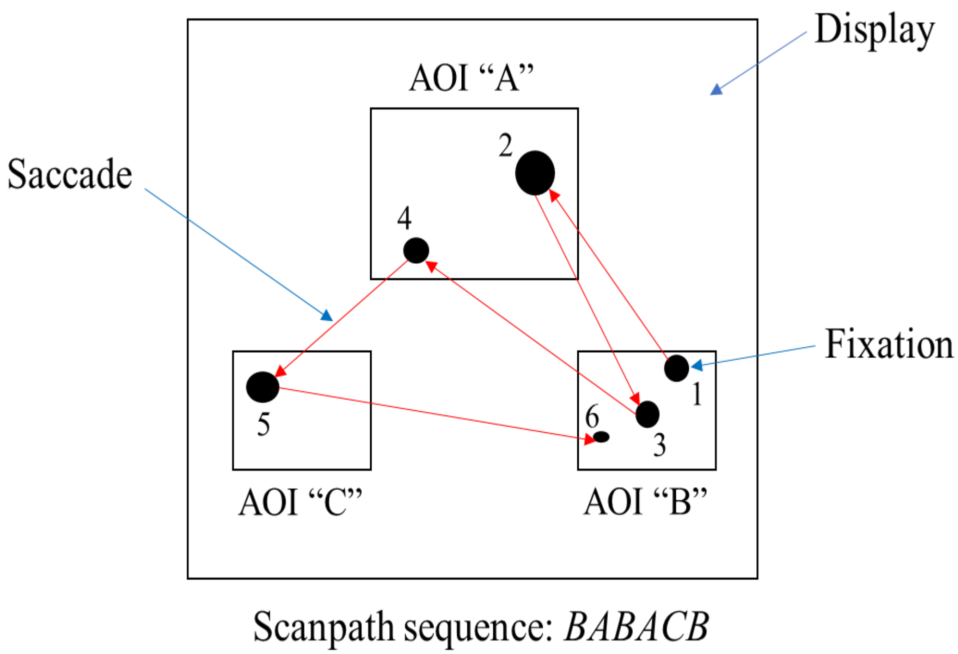
Harrison et al. (2013) defines user satisfaction as “the perceived level of comfort and pleasantness afforded to the user through the use of a system”. Han et al., (2004) defines user satisfaction as how a user feels overall about the interaction with a particular system. Similar to the efficiency attribute, the user satisfaction attribute is frequently used by usability researchers, which provides qualitative insights about a particular system. There are several ways to subjectively evaluate user satisfaction. However, a questionnaire Likert rating scale is perhaps the most effective and commonly used metric to assess the user’s attitude toward a system.

Though Nielsen (1994a) believed that all the aforementioned attributes of the observational method are necessary for informing the usability of interfaces (Nielsen, 1994a), the International Organization for Standardization (ISO 9241-11) (1997) later suggested combining those attributes and grouping them under a model of three main attributes: effectiveness, efficiency, and satisfaction. This later model is supported by numerous scholars and applied extensively in the usability research (e.g. Joo, 2010; Hussain and Kutar, 2009; Georgsson and Staggers, 2015).

Focus groups refer to a group of participants that gather in one place to discuss problems, needs, and goals in a semi-informal setting (Nielsen, 1997). The qualitative findings from focus groups are believed to elicit great knowledge regarding the discussed topic. Details about the focus groups method are presented in Chapter 4.

2.2.3 Eye Tracking-Based Usability Evaluation Technique

Even though the traditional usability evaluation methods provide valuable information to the usability of interfaces, they only give information about the overall performance/experience on the given tasks; they do not give a deep and detailed understanding of the users' interaction with interfaces throughout the entire exposure (Pretorius et al., 2005). To account for this issue, the physiological measurement tool, eye tracking, can be used. Eye tracking gives deep insight towards the usability of interfaces as it provides us with a microscopic view of the user's cognitive processing activity and interfaces design issues (Goldberg & Wichansky, 2003; Fu, 2016). Poole & Ball (2006) defines eye tracking as "a technique whereby an individual's eye movements are measured, so that the researcher knows both where a person is looking at any given time and the sequence in which the person's eyes are shifting from one location to another." Eye fixation occurs when a person directs their visual gaze towards a particular location [two-dimensional coordinate points (horizontal and vertical)] on a display, see Figure 1 (a). In addition, in eye tracking analyses, a display is preferably divided into pre-defined subareas called areas of interest (AOIs). The AOI technique refers to drawing simple geometrical shapes (e.g. rectangles and squares) to represent specific areas of a display that grab the attention of a person through their eye fixations at any stage of the experiment, which enables the analyst to attain statistics on each AOI that reveal meaningful conclusions (Poole et al., 2005), see Figure 1 (a).



(a)

		To		
		A	B	C
From	A	0	1	1
	B	2	0	0
	C	0	1	0

(b)

Figure 1. (a) Simple example of scanpath structured by 6 eye fixations (size of eye fixations is proportional to the duration spent on that fixation) and saccades (transitions between fixations), mapped into 3 AOIs. (b) A transition matrix derived from (a)

There are several eye tracking metrics that can give a clear idea about the interfaces' usability and the users' characteristics. Eye fixation durations and eye fixation frequencies (numbers) per each AOI are among the most informative metrics of this method in assessing the usability (Cheng, 2011; Cooke, 2006). Eye fixation duration refers to the amount of time a person fixates their eyes on a particular object in an AOI (Just & Carpenter, 1976). Several authors believe that fixation duration determines the extent to which the user's cognitive processing is easy or difficult (Graf & Kruger, 1989; Jacob & Karn, 2003). Specifically, Bojko and Schumacher (2008) stated that "longer fixations are a sign of increased difficulty in extracting and processing information due to higher information density, ambiguity, or complexity." Similarly, Cooke (2006) claimed that long eye fixation duration indicates that a user is struggling and/or confused when cognitively processing an element on a display. Eye fixation frequency refers to the number of eye fixations that occur when a user searches for information on AOIs of a display (Goldberg & Kotval, 1999). "Fixation frequency is thought to correspond to search efficiency; the lower the number of fixations on a display, the more efficient the search." (Cooke, 2006). Hence, long eye fixation duration and large eye fixation numbers are indicators of usability problems among system interfaces (Ehmek & Wilson, 2007).

In addition to the metrics of eye fixation durations and eye fixation numbers, the sequential eye movements (scanpath) analysis technique greatly enables us to understand users' thought processes and visual scanning strategies when searching a display for specific information (Ehmke and Wilson, 2007). Specifically, as shown in

Figure 1(a), “scanpath” refers to the route that fixations and saccades take on the display (Noton & Stark, 1971a, 1971b). Goldberg and Wichansky (2002) and Goldberg and Kotval (1999) believe that long scanpath length and duration are associated with less efficient searching and scanning, suggesting inherent usability problems. In addition, Ehmke and Wilson (2007) indicated that scanpaths analysis can greatly help systems’ interface designers optimize their systems based on whether users follow the anticipated scanpath associated with a specific task on a display. One way to understand and quantify this scanpath is by employing the transition matrix approach (Mandal et al., 2016). For example, as shown in Figure 1 (b), the (From) AOIs represent the starting points of the eye fixations, the (To) AOIs represent the destination points of the eye fixations, and the numbers in the matrix represent the number of transitions among AOIs based on the scanpath sequence shown in Figure 1 (a).

2.3 User-Centered Design

Today’s interface designs are not necessarily user-friendly and intuitive to use, leading users to get frustrated, fail to complete a task, and perhaps quit the system and look for alternatives (Abras et al., 2004). An ideal platform example is the smart phone platform. As stated earlier, the smart phone platform, along with its app market, is continuously and rapidly growing among user of all ages. However, a large number of app developers may pay more attention to producing multiple features and highly sophisticated tools than considering the most important factor of the design- the users’ needs (Foraker, 2018). Overlooking user requirements and different characteristics in the design stage may prevent intended users from effectively, efficiently, and/or comfortably using an interface system and achieving its main goals. Systems with poor usability, especially those involving time-critical data, such as weather data, can be associated with several system errors, and/or slow response time, which can hinder performing life-saving actions.

To produce user-friendly app interfaces, developers need to employ a user-centered design (UCD) approach by paying adequate attention to the needs and characteristics of their end-users (Brown et al., 2013).

2.3.1 User-Centered Design Definitions

The term 'user-centered design' was first introduced in the 1980s by Donald Norman and his team at the University of California San Diego (UCSD). However, it has not become popular among researchers until it was published in a book titled: *User-Centered System Design: New Perspectives on Human-Computer Interaction* (Norman & Draper, 1986).

The UCD is defined in several different ways, however, all definitions depict the same idea: designing an interface based on the user requirements. A widely accepted definition is presented by Norman (1988): iterative procedure where designers prioritize user needs in each of the design stages. The ISO 9241-210 (2010) also defines the UCD as a cycle of design phases where interface developers consider intended users' needs, capabilities, and limitations in each design phase. Similarly, Abras et al. (2004) stated that the UCD is "a broad term to describe design processes in which end-users influence how a design takes shape."

2.3.2 User-Centered Design Processes and Evaluation Techniques

Norman (1988) in his book: *The Psychology of Everyday Things (POET)* explained the UCD process and how it should be centered around the user. Specifically, he introduced four guidelines for developers when designing interfaces to make the user at the center of the design. The guidelines are as follows:

- Simplify the system in a way that any action is expected at any moment.
- Enhance the visibility of the system in which each operational step is recognized by the user.
- Simplify the system in a way that the user can easily recognize the status of the system.
- Promote mapping between user input and resulting action in a way that matches with the user's mental model.

Norman (1988) then realized that recommending developers to create intuitive and easy interfaces is vague; specific design guidelines are needed. He developed seven guidelines that are listed as follows:

- Prior to implementing the design, create simple manuals based on prior knowledge.
- Do not load the interface with too much content and several navigational steps, as users have limited short-term and long-term memory.
- Make menus, icons, and all texts visible so that the user can easily figure out the required action.

- Use graphics so that the user can quickly understand the objects' functionality and easily map between the objects' shapes and resulted actions.
- Use constraints, when needed, to inform the user that there is only one option to perform an action.
- Make the system error-free. If not possible, provide clear and actionable error messages and facilitate the user's recovery from the error made.
- If you fail to meet the preceding guidelines or your system must include subjective mapping, create your own international standard.

Based on these design guidelines, Nielsen (1995a) developed a set of usability heuristics for usability engineering that follow the same concept of the guidelines by Norman (1988). Nielsen's heuristics are discussed in detail in Chapter 4.

Even though these guidelines are extremely important for the success of any interface design, it is crucial in the UCD process to involve actual users and learn about their exact needs and limitations with respect to the system of interest. Therefore, a common UCD framework grounded by Norman (2013) and Mao et al. (2005) and widely used in the field of human factors, such as in Schnall et al. (2016) and Witteman et al. (2015), is depicted in Figure 2.

This framework is generated based on the idea that a system or product is expected to meet intended users' needs, if the development process is a cyclical process in which users' needs are considered at each stage. As shown in Figure 2, the UCD process starts with the "User" phase in which the design team collects data from actual users regarding their needs in, concerns about, and limitations with a specific system or

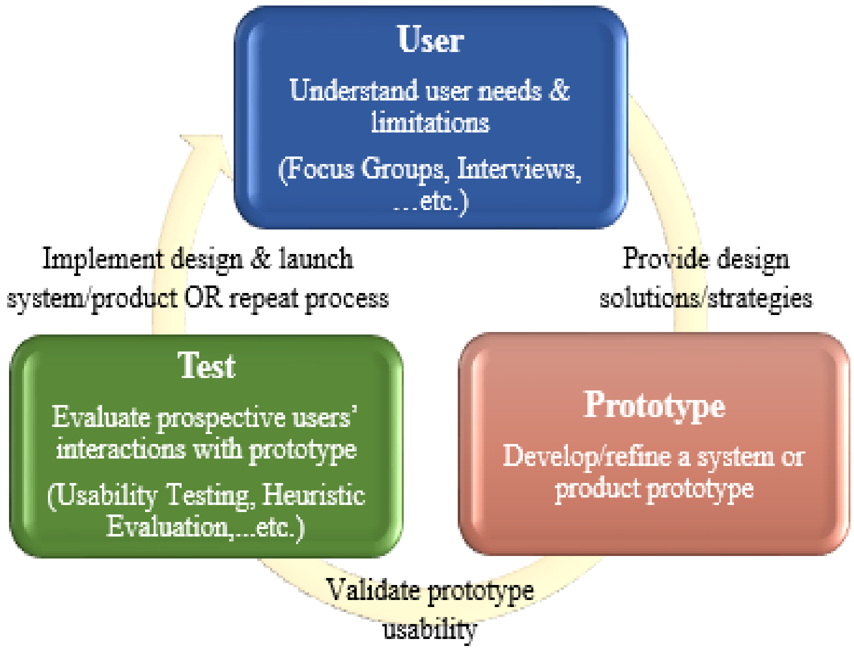


Figure 2. Framework of user-centered design

product. The data in this phase is usually collected qualitatively using one-to-one interviews, focus groups, and personas. Once the data are collected and carefully analyzed, developers create design strategies and solution manuals to meet users' requirements in a way that matches with general design guidelines such as those mentioned earlier. In the next phase "Prototype", developers build/refine the system or product prototype considering the outputs from phase 1. Once the prototype is developed/refined, it is time to validate the prototype usability and insure whether it exactly meets users' needs generated from phase 1. In phase 3 "Test", actual users are involved again by having them interact with the developed prototype based on a set of tasks. The evaluation in this phase is mostly performed using a usability testing method, which is measured quantitatively and/or qualitatively. There are also other methods, such as heuristic evaluation by experts, used for validation. Once phase 3 is completed and the results are analyzed, the designers have the option either to implement the design and launch the system/product to the market or repeat the whole process again, depending on the findings from phase 3.

2.4 Previous Usability and User-Centered Design Research of Smart Phone Weather Applications

Several studies have focused on examining the usability of smart phone apps in different specific areas such as in tourism (Geven et al., 2006; Shrestha 2007; Ahmadi and Kong 2008; Schmiedl et al., 2009) and geography (Elzakker et al., 2008). However, very few studies directly pertained to weather alert apps. All those studies employed traditional metrics for evaluating the usability. Throughout the evaluation on this area, two consecutive studies by Singhal (2011) and Alluri (2012) examined the usability issues in the interface design of all the originally built-in smart phone apps in iPhone and Android, respectively. The researchers investigated users' (three users in Singhal's (2011) study and five users in Alluri's (2012) study) general understanding of symbols and icons, speed of performing common tasks, and the ease of using the apps in general. Both studies revealed a few issues in each app such as lack of visibility, lack of affordance, and poor consistency. For example, the lack of visibility was present in the weather app in both studies, where participants could not easily see the weather information icon "i" because of its very small size. However, as both studies tested the native (originally built-in) weather apps that included only basic weather forecasts, the findings may not be sufficient and/or compliant with the actual demand in the field.

This is said due to the high and increasing dependency of users on accessing detailed weather information with highly sophisticated features through downloadable weather apps, as stated earlier.

The most applicable usability evaluation study of smart phone weather alert apps to the work in this dissertation was conducted by Drogalis et al. (2015). By recruiting six participants, the researchers evaluated the performance of participants on several tasks included under three main features in the "Weather Channel" app, in terms of task completion time, Likert ratings of the tasks, and comments made by participants. These features were weather and location settings, iWitness weather account, and pollen alerts. Even though multiple usability issues were determined from the subjective evaluation metrics used in Drogalis et al.'s (2015) study, the completion time of the given tasks did not provide adequate judgement of the users' performance. For example, the results showed that participants took an average completion time of five minutes to create an iWitness account, while the other tasks did not exceed one minute and thirty seconds on average. The study did not consider a benchmark approach that links the tasks' completion time recorded from the users to a standard data in order to logically determine whether the user's performance was satisfactory or poor. Instead, the researchers listed the completion times of all the tasks and arbitrarily concluded that one of the tasks yielded a long completion time, while the other tasks had short completion times. In addition, even though usability issues can be determined from only a limited number of users (3-5 users), from a statistical standpoint, at least twenty participants should be involved (Nielsen, 2012). Drogalis et al.'s (2015), Singhal's (2011), and Alluri's (2012) studies included only six, three, and five participants, respectively.

As noted earlier, eye tracking is considered one of the most advanced analytical tools for users' cognitive processes and decision-making strategies. Scholars recommend incorporating this tool with traditional tools in usability evaluation in order to uncover more usability issues as well as determine the reasons for those issues in an objective manner (Pretorius et al., 2005). The eye tracking assessment tool has been utilized in many different fields to evaluate the usability of different systems' interface displays, such as in neural network on tablets (Holland et al., 2013) and websites of educational multimedia on desktop computers (Yen and Esgin, 2015). However, the use of eye tracking in evaluating the usability of apps in smart phone devices in general and in weather alert apps in particular is still significantly lacking. Among the very few studies using eye tracking regarding the usability evaluation of smart phone apps is a study performed by Chynal et al. (2012). In their study, they compared users' performances on the Facebook app shown on two different display types: a smartphone and a personal computer (PC) emulator using conventional (i.e. completion time and surveys) and eye tracking metrics. The study attained interesting results from both the conventional metrics and the eye tracking metrics. However, for the eye tracking part, the smartphone display was treated as only one overall AOI and no specific AOIs were considered. In the eye tracking assessment section of this dissertation, we employ the AOIs technique by dividing the app's display into multiple AOIs in order to get a comprehensive and meaningful information about the user's cognitive and decision-making strategies when scanning the display elements.

Similar to the usability evaluation of smart phone apps in the weather domain with the use of eye tracking, to the best of our knowledge, no work is published using the UCD approach on smart phone weather apps. In addition, even though an influx of research has focused on age differences when using smart phone apps in various domains, including information technologies (García-Peñalvo et al., 2014), healthcare (Cáliz and Alaman, 2014; Morey et al., 2017), and communication (Smith and Chaparro, 2015; Al-khomsan et al., 2015), a UCD approach that involves the needs and limitations of different age groups in weather apps' design process is still lacking for research. In particular, little consideration has been devoted for older users' special characteristics in smart phone apps design. This is supported by the findings from several studies (e.g. Fisk et al., 2009; Lorenz and Oppermann, 2008). Older users' characteristics, limitations and requirements are explained in Chapter 4.

However, a relevant UCD study with the focus on end-users' requirements in smart phone apps was done by Liu (2012). Specifically, Liu (2012) qualitatively examined the perceived usability (PU) and the perceived ease of use (PEU) of two age groups on four different smart phone apps. The two groups included younger users (ages 18 – 30 years old) and older users (50+ years old). The results revealed different usability scores between the younger and older users. More specifically, younger users had higher scores on the PU and the PEU than older users. In addition, the researcher assessed the differences between the two groups in terms of 12 usability characteristics such as font size, function keys, and scrolling menu. The results showed that the younger group ranked the most important usability characteristics of the given smart phone apps differently than older users. In particular, the most important usability characteristics for the younger users were “intuitive menu options,” “prevention of making errors,” “appropriate font size,” and “minimum number of steps to accomplish tasks”. On the other hand, older users reported three different important usability characteristics: “appropriate number of function keys,” “ease of prompts and cues,” and “effortless menu scrolling” and they agreed with younger users on only one characteristic: “appropriate font size”. These findings clearly imply that younger and older users’ needs might be different even though the aforementioned characteristics are important to be considered in all interfaces and for all user and age groups.

2.5 Chapter Summary

This chapter has presented a review of the work cited in the literature regarding smart phone technology with a particular focus on the smart phone app market, usability evaluation of smart phone weather apps using different evaluation methods, and UCD assessment for different age groups. Overall, very few studies were identified to directly pertain to the usability and user-centered designs of smart phone weather apps. Furthermore, several gaps and limitations were recognized and will be addressed in this dissertation. The gaps and limitations include: 1) insufficient sample size, as in Drogalis et al.’s (2015), Alluri’s (2012), and Singhal’s (2011) studies; 2) insufficient knowledge acquired from the tested smart phone apps, as in Alluri’s (2012), and Singhal’s (2011) studies; 3) lack of a benchmark approach or a standard reference measure, as in

Drogalis et al.'s (2015) study; 4) lack of objective assessment tools, such as eye tracking, that examine user's thinking processes and decision-making strategies in smart phone weather apps; 5) inappropriate utilization of AOIs, as in Chynal et al.'s (2012) study; and 6) lack of a comprehensive UCD research of smart phone weather apps that considers users' goals, needs, issues, and characteristics of different age groups throughout the entire UCD process. Conducting research in these areas will create a solid research foundation, as well as help users navigate user-friendly interfaces of real-time-critical information, such as weather interfaces. In addition, businesses are expected to experience success and growth as long as they design interfaces with users' needs and characteristics in the center of the development process.

Chapter 3. A Mixed Methods Approach to Evaluating the Usability of Smart Phone Weather Alert Applications

Author's Note: The content in Section 3.2 was published as a journal paper, titled "Usability evaluation of mobile weather hazard alert applications", in the Industrial and Systems Engineering Review Journal. The content in Section 3.3 was submitted as a journal paper, titled "Integrated eye movement analysis approach of time-critical services: application in a mobile weather alert system", to the Industrial Engineering and Management Systems Journal. The author of this dissertation wrote both papers in collaboration with his advisor, Dr. Ziho Kang.

3.1 Introduction

In this chapter, we present a mixed methods approach to determining the usability problems that result from users' interaction with smart phone weather apps and understanding users' thinking processes when performing given tasks. Specifically, we present two studies for the evaluation. In study I, we use traditional usability evaluation methods using the ISO 9241-11 (1998) model: effectiveness, efficiency of use, and user satisfaction. In study II, we use an eye tracking-based method with the aim to examine whether the eye tracking can better support the analysis of such apps and predict users' performance and cognitive processes from their eye movements.

As stated earlier that usability problems with interfaces are mostly uncovered by novices, both studies in this chapter mainly focused on this population. However, to address the limitation in Drogalis et al.'s (2015) study of not having a reference measure for evaluation, we included experienced users in both studies for that purpose only. Both studies used one of the most popular and widely used weather apps, Weather Radio, as a sample representative of all weather apps. The Weather Radio app is created and run by Weather Decision Technologies (WDT) company (Weather Decision

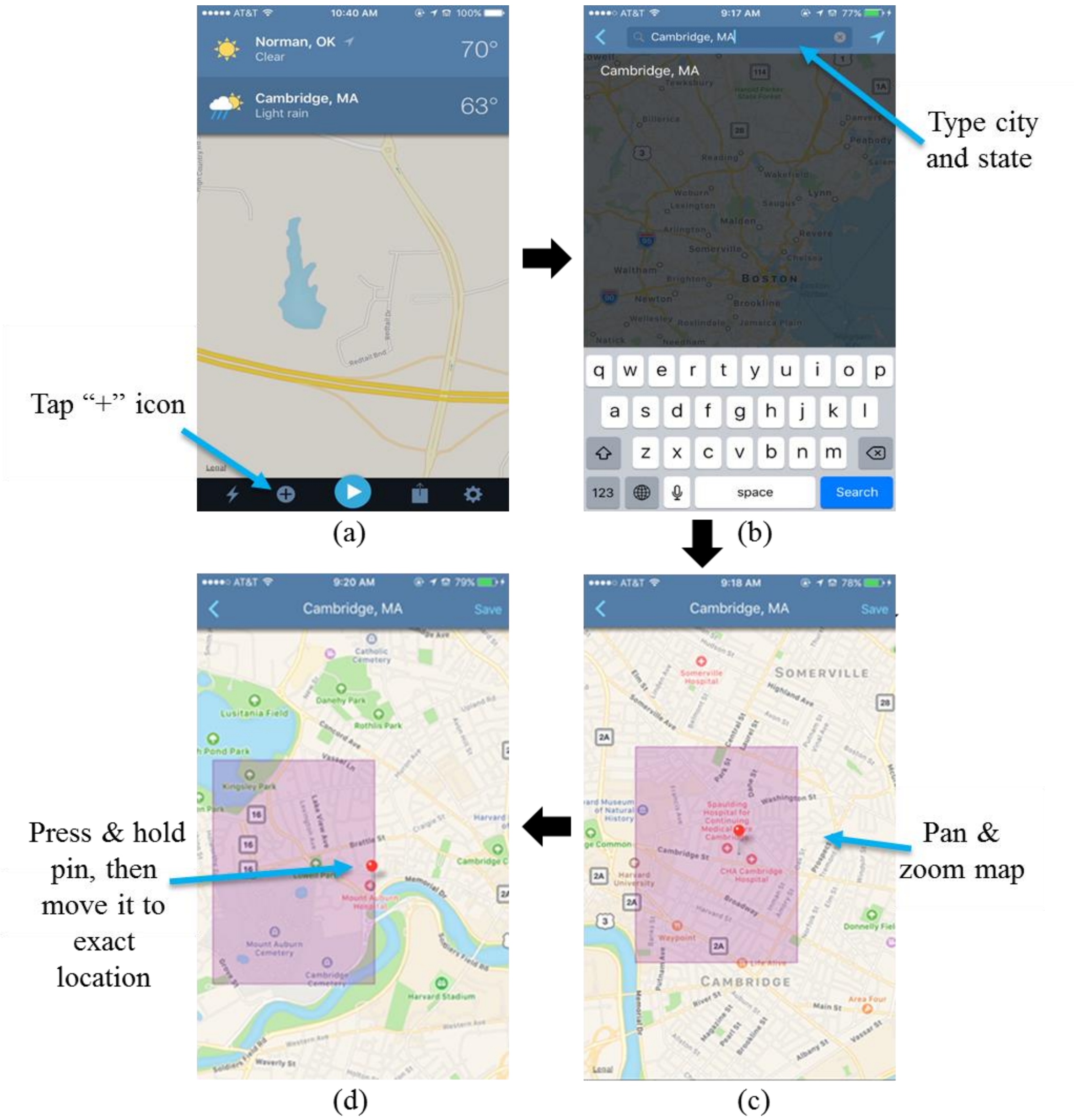


Figure 3. Example of location search process on Weather Radio

Technology, 2018). It is worth mentioning that almost all popular weather apps have similar features: location search, weather forecasts, alert messages, map settings, and alert settings; the difference might exist in how they present information. All of Weather Radio app’s features are explained in text and graphs as follows.

- **Location Search**

Figure 3 shows the process of using the location search feature on Weather Radio. Unlike other weather apps that use only zip code and/or city & state, Weather Radio enables users to add an exact location in addition to traditional search methods. To add an exact location, a user has to (a) tap the "+" icon at the bottom of the navigation menu, (b) type the city and state within the app's search bar and select it once it appears from the auto-suggestions list, (c) navigate and zoom in/out the map of the selected location until finding the desired exact location, and (d) press and hold the red pin until lifted to move it and drop it at the exact location; then save it from the top right screen.

- **Weather Forecasts**

Figure 4 shows the process of accessing specific weather forecasts on Weather Radio. To access weather forecasts of any desired saved locations, a user needs to (a) tap the location, then (b) will appear. To access current extended weather forecasts (e.g. humidity & wind) as shown in (c), the top part in (b) needs to be tapped. To access hourly temperature forecasts as shown in (d), the corresponding day in (b) needs to be tapped. To access weather forecasts of other saved locations, a user needs to go back, by tapping the top left arrow, to (a) and repeat the process.

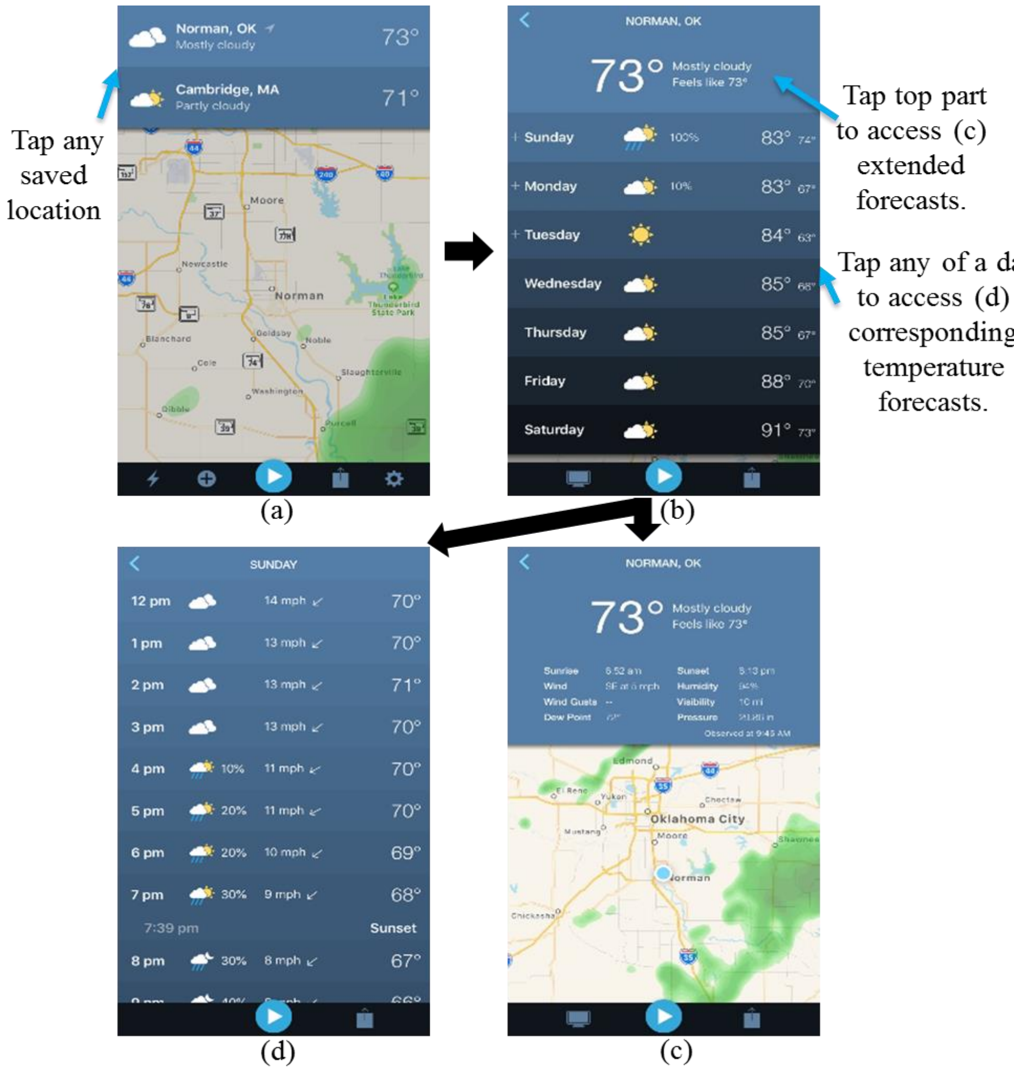


Figure 4. Example of accessing weather forecasts process on Weather Radio

- **Alert Messages**

Figure 5 shows the process of accessing specific detailed messages on Weather Radio as soon as the corresponding alerts are issued by local weather agencies. To access the pushed alert message as shown in (b), a user needs to tap the white alert icon associated with the corresponding affected location as shown in (a). To access the whole alert message content in (b), a user needs to scroll the black box up and down.

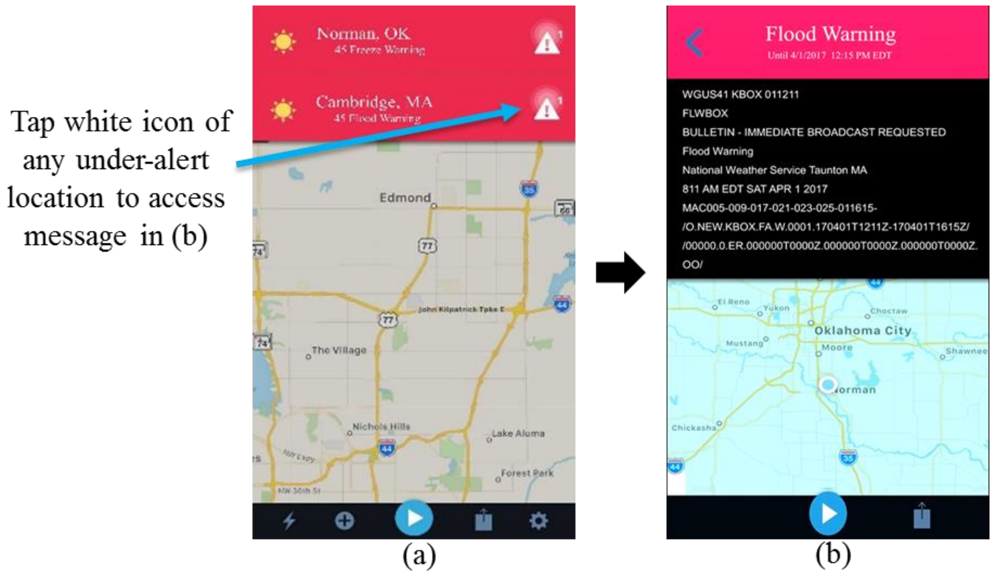


Figure 5. Example of accessing alert messages process on Weather Radio

- **Alert Settings**

Figure 6 shows the process of controlling alert settings on Weather Radio. To control alert settings, a user needs to (a) tap the gear icon at the bottom right of the navigation menu, (b) tap “NWS Alert” from the general settings menu, (c) tap any desired main alert, and (d) enable or disable the listed sub-alerts. Both alerts and sub-alerts lists need to be scrolled up/down to access the whole list.

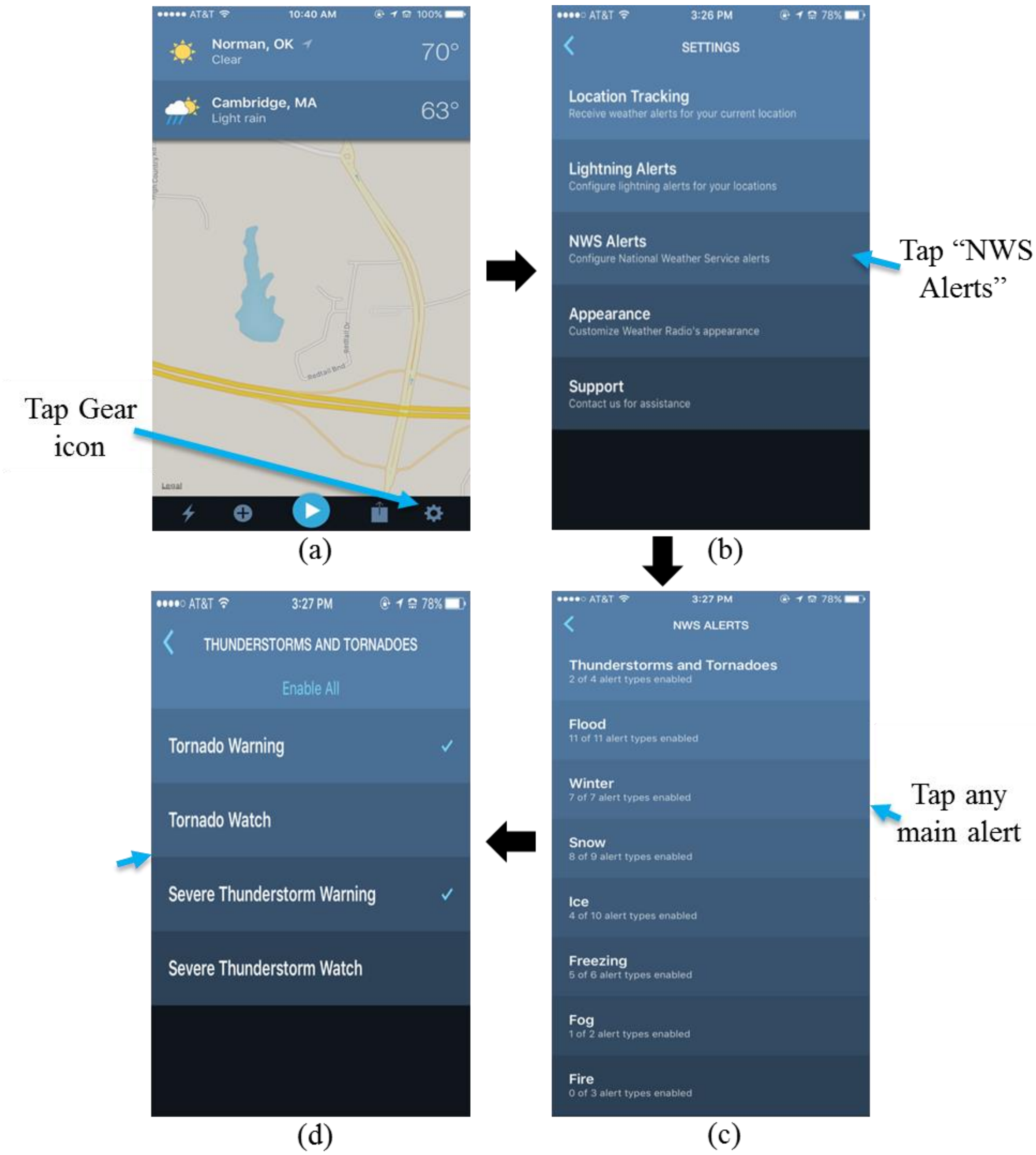


Figure 6. Example of controlling alert settings process on Weather Radio

- **Radar Maps**

Figure 7 shows the process of accessing maps of saved locations and controlling map settings on Weather Radio. To view a particular location's map and control its settings, a user needs to (a) tap the desired location, (b) scroll up the screen and tap the

Figure 7. Example of controlling radar map settings process on Weather Radio bottom map portion to see the radar, (c) tap any icon at the bottom navigation menu to control the corresponding map settings, and (d) access the corresponding settings and change them to their preferences.

3.2 Study I: Traditional Usability Analysis of Smart Phone Weather Applications

3.2.1 Objective

This study aims to determine the usability problems that result from first-time users as well as propose approaches to test to what extent the proposed approaches help enhance the usability of the weather alert apps in comparison to the existing features. Specifically, this study focuses on investigating four important features in the Weather Radio app: (1) searching for locations using two methods: dragging the pin on the map and typing the address in the app's text bar; (2) changing the alert settings; (3) changing the map settings; and (4) comparing two sets of weather alert messages: existing NWS alert messages vs. proposed alert messages. In addition to the given tasks, an exit survey is given to participants asking about their experience during the experiment, as well as about the overall usability of the Weather Radio app.

3.2.2 Method

Participants

A total of 40 participants (users) were recruited for the experiment. All participants were students from the University of Oklahoma (OU), Norman Campus and were regular smart phone (iPhone) users at the time of the experiment. The users were randomly divided into two groups: 1) 20 users with comprehensive training on the Weather Radio app (experienced users) and 2) 20 first time users. One of the study researchers provided the training sessions to the experienced users group. The age of users ranged from 21 to 44 years (Mean (*M*) = 24.70, Standard Deviation (*SD*) = 4.89 years). Both the first-time and experienced users performed all the given tasks. Even though the usability issues were mostly determined from first time users' interaction with interfaces, experienced users were included in the experiment in order to provide standard data for comparative evaluation and add more insight to the current usability of the weather apps.

Apparatus

The Weather Radio (version: 3.0.5) (http://weatherradioapp.com/) was installed and run on a smartphone (iPhone 6). A stopwatch was used to collect the response times for each of the given tasks. The demographic survey, the different types of alert messages, and the exit survey were printed out on paper.

Procedure

In a laboratory setting, users were first provided with an informed consent form. Upon agreeing to participate in the study, users were given a short survey asking about some demographic information. Next, half of the users (20 users) received comprehensive training on the Weather Radio app's features. In addition, they were given time to practice navigating the app's interface by themselves and to ask questions if needed; they were asked to verbally state, "I am ready to begin the experiment", once they felt comfortable with the app. The average time of the training sessions from the beginning until users stated they were ready to begin the experiment was 8.43 minutes. The other half of the users (20 users) were completely new to the app and received no training at all. Following that, the users were informed that the experiment would include four tasks: location search, alert settings, and map settings tasks to be completed using the smart phone device, and the alert messages task to be completed by pen and paper. Then, the experiment began, and the tasks were counterbalanced across participants. Tasks instructions were given to participants on a sheet of paper. One of the study researchers observed the participants' interaction with the tasks performed on the smart phone device by recording the tasks' completion time.

The weather alert messages task was not accomplished on the Weather Radio app because the original weather alert messages only appear when there is a weather alert in effect at that time. They were recorded prior to the experiment and then compared by all users with the proposed messages (see details in the corresponding sub-section). In addition, our main goal with this task was to assess user comprehension and satisfaction when reading existing NWS messages and whether the proposed modifications would enhance their experiences. At the end of the experiment, all users completed an exit survey to evaluate their experience with all the given tasks, as well as their opinions toward the overall usability of the app (see Appendix A for exit survey questions).

Tasks

- **Location Search Task**

The location search task was to find a specific location using two approaches: pin icon allocation and typing. The pin icon allocation approach was a feature implemented by the Weather Radio app; see section (3.1) for details about the process of performing this task). The purpose of this feature was to search the embedded Google map for a specific location for which a user can access weather forecasts. This pin approach is utilized by moving the pin icon on the map to the location of interest. The typing approach was to type the local address on the text bar instead of having to move the pin icon. This approach was not an active feature in the Weather Radio app, but was included by the study researchers in order to compare it with the pin icon approach and then determine which approach would be more efficient. Specifically, for the location search task, we assumed that the family member of an end-user, the role played by the test participant, is at Mount Auburn Hospital in Cambridge, MA.

For the pin icon allocation approach, the task instruction given to participants was as follows: *"Please find the Mount Auburn Hospital in Cambridge, MA using the pin icon on the embedded Google map."* For the typing approach, the task instruction given to participants was as follows: *"Please find the Mount Auburn Hospital in Cambridge, MA by typing (330 Mount Auburn St., Cambridge, MA 02138) in the app's text bar."*

- **Alert Settings Task**

The alert settings task was to change settings of certain weather alert notifications. In particular, the participants were asked the following: *"Please enable (turn on) Tornado Warning and Severe Thunderstorm Warning and disable (turn off) Tornado Watch and Severe Thunderstorm Watch."*; see section (3.1) for details about the process of performing this task.

- **Map Settings Task**

The map settings task was to change the settings of the map type and the weather layer. More specifically, the participants were asked the following: *"Please change the map type from Standard to Hybrid and the weather layer from Radar to Clouds."*; see section (3.1) for details about the process of performing this task.

- **Weather Alert Message Evaluation Task**

This task included two examples of weather alert messages that previously appeared on the Weather Radio app to alert users about current and future weather threats. The first weather message, Severe Thunderstorm Watch (STW), appeared on the app on March 30, 2016, to warn users about a severe thunderstorm watch; and the second one, wind advisory (WA), appeared on Mar 21, 2016, to inform users about a wind advisory. Each weather message was compared as a sample with its proposed message based on statements with a Likert rating scale from 1 to 10, where 1 stands for 'strongly disagree' and 10 means 'strongly agree'. Higher rating scores mean positive opinion and lower rating scores mean negative opinion.

The original version of the STW message was compared with the proposed version of the STW message (see Appendix B (1) for entire original and proposed STW messages). Similarly, the original version of the WA message was compared with the proposed version of the WA message (see Appendix B (2) for entire original and proposed WA messages). The experiment's researchers created the proposed messages. Both proposed messages had the same content as the original messages, except contextual information related to usability was included in the proposed messages. The contextual information refers to additional and interpretive information and language tools that explain unfamiliar words, codes, and symbols in ways that are easy to understand. Examples of the contextual information applied in the proposed messages included using appropriate delimiters (i.e. punctuation marks), upper-case and lower-case letters, easy and intuitive terminology, hierarchical structure based on priority, and comprehensive expressions. Applying such information is believed to enhance the users' overall comprehension of the alert messages as well as the quick physical and mental reaction to the potential weather threat included in the message.

- **STW Messages**

The original and proposed versions of the STW messages had four pairs of statements. The first pair inquired about the understandability of the header information in each message with the presence of the definitions and meanings of weather terms in the proposed message and the absence of the definitions and meanings in the original message. The second pair asked about the readability and understandability of the format of the information about areas under alert using no delimiters in the original message, while using delimiters in the proposed message. The third pair wondered about the readability and understandability of the format of the messages' information using only upper-case letters in the original message and using both upper-case and lower-case letters in the proposed message. The last pair of statements was about the extent to which users were satisfied with the content and organization of both messages (see Table 1 for more details).

- **WA Messages**

Similarly, the original and proposed WA messages had four pairs of statements. The first pair was about the appropriateness of the location of the WA information and the expected impact information of the WA in both messages. The WA information and the expected impact information were located at the end of the original message, while they were located at the top in the proposed message. The second pair of statements was

Table 1. Survey statements for the original and proposed STW messages

	Original NWS message statements	Proposed message statements
Header	**1)** I believe that the header information in this message significantly helped me to understand the alert message: "WATCH COUNTY NOTIFICATION FOR WATCH 58 NATIONAL WEATHER SERVICE NORMAN OK 150 PM CDT WED MAR 30 2016 OKC015-017-019-027-031-033-047-049-051-053-067-071-073-081-083-085-087-099-103-109-119-125-137-TXC009-023-077-485-310200- /O.NEW.KOUN.SV. A.0058.160330T1850Z-160331T0200Z./"	**1)** I believe that the header information in this message significantly helped me to understand the alert message: "Severe Thunderstorm Watch # 58 in 2016: For Counties of Oklahoma, Counties of Texas, and Cities that include the impacted counties. Time: 1:50 PM, Time Zone: Central Daylight Time (CDT), Day: Wednesday, Date: 03/30/2016."
Use of Delimiter	2) I find using "…" for separation between the areas under alert marks significantly enhanced the readability and understanding of this message. For example: "THIS INCLUDES THE CITIES OF… ANADAEKO…ARCHER CITY ARDMORE…BLACKWELL…BLANC HARD…CHANDLER…CHICKASHA …CONCHO…DAVENPORT… DAVIS…DUNCAN…ELRENO…ENID …GUTHRIE…HENNESSEY… HENRIETTA…HINTON...HOLLIDAY …KINGFISHER	**2)** I find using some punctuation marks (":", "-") for separation between the areas under alert significantly enhanced the readability and understanding of this message. For example: "Counties: OK: Central: Cleveland - Grady - Canadian - Kingfisher - Lincoln - Logan - McClain Oklahoma - Payne - Pottawatomie Northern: Kay - Garfield - Grant - Noble Southern: Carter - Jefferson - Garvin - Love - Murray."
Letters Format	**3)** I find using only upper-case letters significantly enhanced the readability of this message.	**3)** I find using both upper-case and lower-case letters significantly enhanced the readability of this message.
Satisfaction	**4)** Overall, I am satisfied with the content and organization of this message.	**4)** Overall, I am satisfied with the content and organization of this message.

about the comprehensive word expressions of the wind information by using technical expressions and concepts in the original message and by using equivalent everyday life examples in the proposed message. The third pair was about the use of terminology using jargon in the original message and using common terminology in the proposed message. The last pair of statements was about the extent to which users were satisfied with the content and organization of both messages (see Table 2 for more details).

Table 2. Survey statements for the original and proposed WA messages

	Original NWS message statements	**Proposed message statements**
Information Location	**1)** I believe that the Wind Advisory information including the Impacts information located **at the end of this message** is appropriate. "...WIND ADVISORY REMAINS IN EFFECT FROM 11 AM TO 9 PM CDT TUESDAY... * TIMING...11 AM TO 9 PM. * WINDS...SOUTH TO SOUTHWEST 25 TO 35 MPH WITH GUSTS 40 TO 50 MPH. * IMPACTS...DRIVING COULD BECOME DIFFICULT ESPECIALLY IN HIGH PROFILE VEHICLES. ANY LOOSE OUTDOOR ITEMS COULD ALSO BLOW AROUND."	**1)** I believe that the Wind Advisory information including the Impacts information located at the top of this message is appropriate. "Wind Advisory: For Counties of Oklahoma. Time: 4:56 PM, Time Zone: Central Daylight Time (CDT), Day: Monday, Date: 03/21/2016 Wind Advisory remains in effect from 11 AM CDT on Monday to 9 PM CDT on Tuesday. Impacts: Driving could become difficult especially in tall vehicles. Any loose outdoor items could also blow around. Avoid riding motorcycles or bicycles."
Word Expressions	**2)** I find using wind speed information such as "SOUTH TO SOUTHWEST 25 TO 35 MPH WITH GUSTS 40 TO 50 MPH" more useful than using equivalent alert messages of the wind impact using real life examples such as "Avoid riding motorcycles."	**2)** I find using alert messages about the wind impact using real life examples such as "Avoid riding motorcycles" more useful than using wind speed information such as "SOUTH TO SOUTHWEST 25 TO 35 MPH WITH GUSTS 40 TO 50 MPH."
Terminology	**3)** I find the terminology used in this message completely understandable such as the bolded phrase in this quoted text "DRIVING COULD BECOME DIFFICULT ESPECIALLY IN **HIGH PROFILE VEHICLES.**"	**3)** I find the terminology used in this message completely understandable such as the bolded phrase in this quoted text "Driving could become difficult especially in SUVs or trucks"
Satisfaction	4) Overall, I am satisfied with the content and organization of this message.	**4)** Overall, I am satisfied with the content and organization of this message.

Variables

The study included two independent variables associated with the given tasks. Each independent variable had two levels. For the location search and weather alert messages tasks, both independent variables were included: experience and type of approach. The levels of experience were first-time and experienced users. The types of approach were a pin on a map and typing in the app's text bar for the location search task, while they were original NWS and proposed messages for the weather alert messages task. In addition, the alert settings and map settings had one independent variable: experience with the same levels as in location search and alert messages tasks.

Three dependent variables were included in this study: task completion rate, task completion time and survey Likert rating score. The task completion rate was to determine the effectiveness of the app's interface based on the three search tasks: location search, alert settings, and map settings. The task completion time was used to assess the users' efficiency on the three search tasks. The survey Likert rating score was used for the alert messages to examine how users subjectively evaluate and compare between the content and format of original and proposed alert messages.

Data Analysis

A Two-Way Mixed Design ANOVA or independent sample t-test was used for the experiment's tasks. Specifically, the Mixed Design ANOVA test was used for the location search and alert messages tasks. The independent sample t-test was used for both the map settings and the alert settings tasks to compare the data collected from the first-time users with the data collected from the experienced users. Finally, for the exit user satisfaction survey, descriptive statistics analysis or qualitative content analysis was used.

3.2.3 Results

Effectiveness

The effectiveness results showed that all users were able to successfully complete the given tasks (100% task completion rate).

Efficiency

- **Location Search**

Figure 8 shows graph comparisons between the two approaches towards the location search task as well as between the two user groups in terms of the mean completion time with Standard Error (SE) bars.

A Two-Way Mixed Design AVOVA with approach type (pin on map and app's text bar) as a within-subjects factor and user group (first-time and experienced users) as a between-subjects factor was conducted, in terms of the task completion time. The results showed that there was a significant main effect of approach type (F (1, 38) = 49.13, $p < .001$) on the task completion time, with significantly shorter completion time on the app's text bar approach (M = 48s), compared to that on the pin on map approach (mean = 139.33s). On the other hand, there was no significant main effect of user group (F (1, 38) = 1.16, $p = .289$) on the task completion time, as first-time (M = 100.93s) and experienced (M = 86.40s) users performed similarly. In addition, there was no significant interaction effect between approach type and user group (F (1, 38) = .34, p = .562), meaning that the task completion time observed on each location search approach does not depend on the user's experience level.

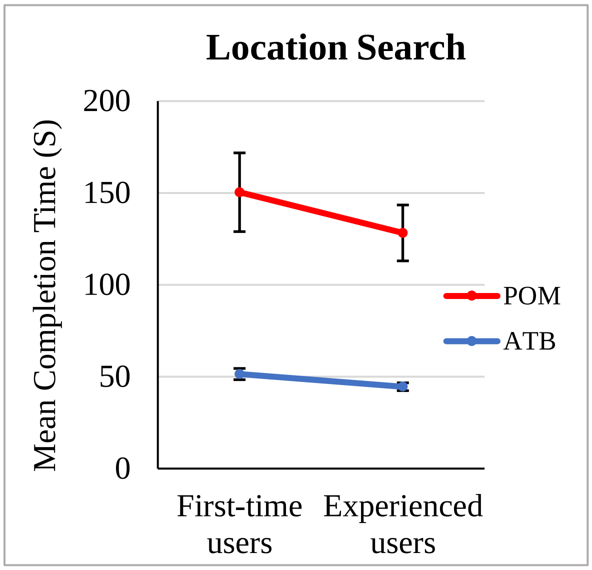


Figure 8. Plot of task completion time for the location search; POM: pin on map and ATB: app's text bar

- **Alert Settings**

Figure 9 shows graph comparison between first-time and experienced users on the alert settings task in terms of the mean completion time with SE bars.

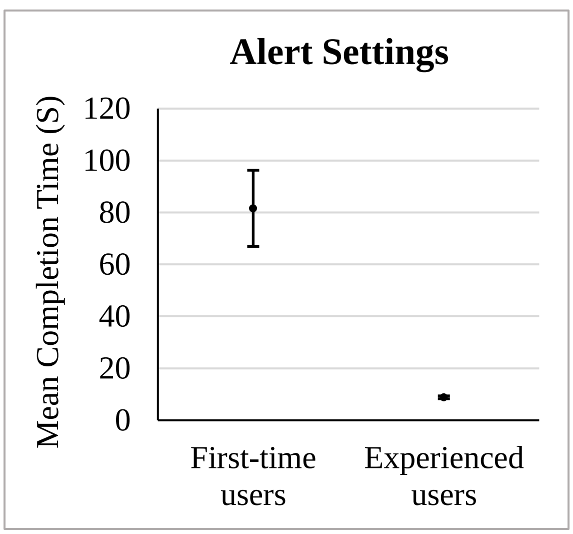


Figure 9. Plot of task completion time for the alert settings

An independent sample t test was performed to determine whether a difference existed between the first-time users and the experienced users, in terms of the mean completion time on the alert settings task. The results revealed a statistically significant difference, t (38) = 4.960, $p < .001$, indicating that the mean completion time of the experienced users (M = 8.85s, SD = 2.52s, N = 20) was significantly less than the mean completion time of the first-time users (M = 81.60s, SD = 65.55s, N = 20).

- **Map Settings**

Figure 10 shows graph comparison between first-time and experienced users on the map settings task in terms of the mean completion time with SE bars.

Similar to the alert settings task, an independent sample t test was performed to determine whether a difference existed between the first-time users and the experienced users, in terms of the mean completion time on the map settings task. The results revealed a statistically significant difference, t (38) = 8.459, $p < .001$, indicating that the

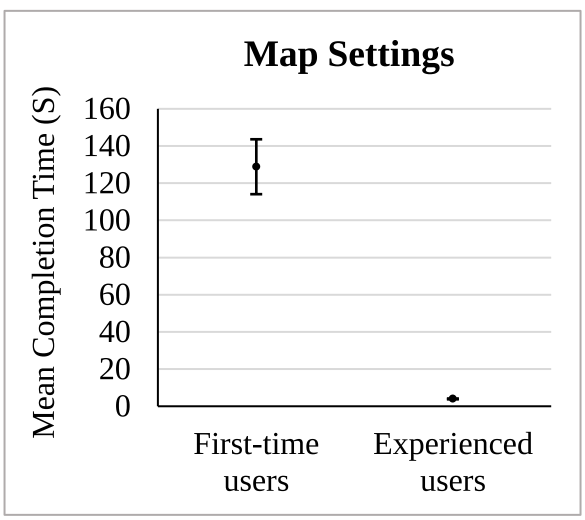
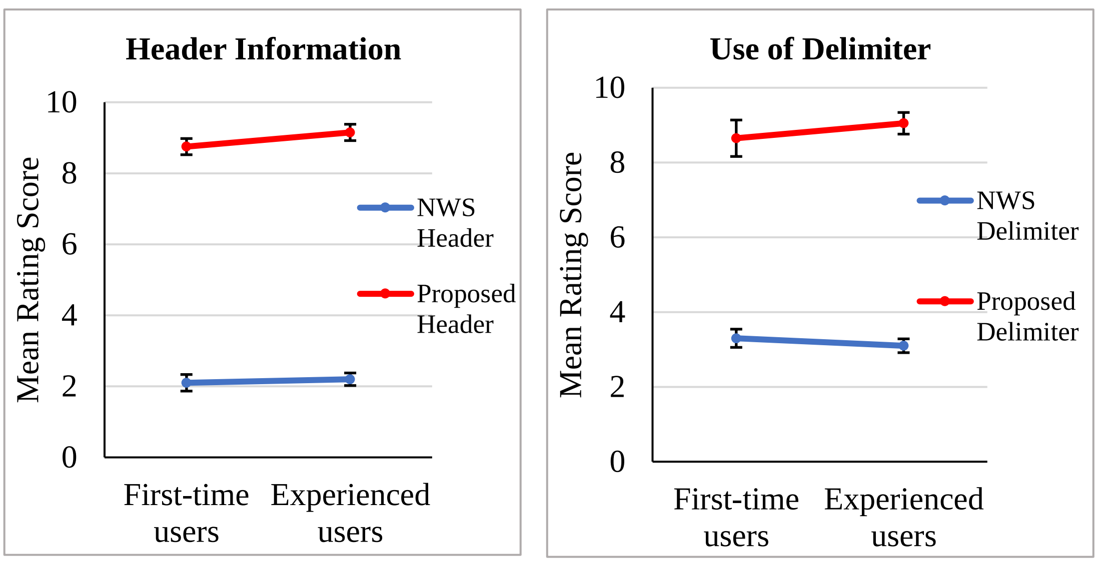


Figure 10. Plots of task completion time for the map settings

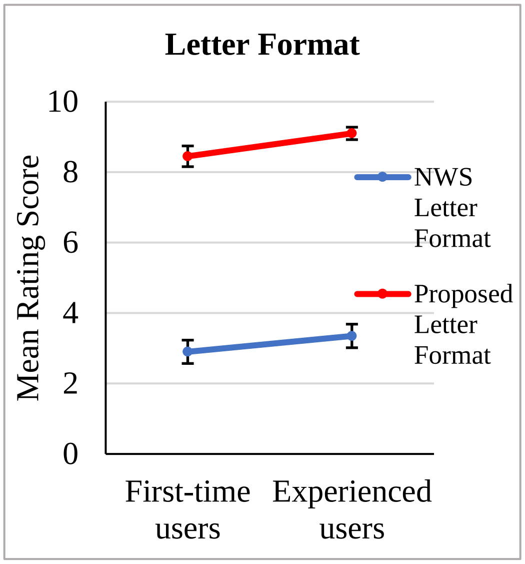
mean completion time of the experienced users (*M* = 3.95s, *SD* = .94s, *N* = 20) was significantly less than the mean completion time of the first-time users (*M* = 128.85s, *SD* = 66.03s, *N* = 20).

- **Survey Comparison of Weather Alert Messages: STW alert messages**

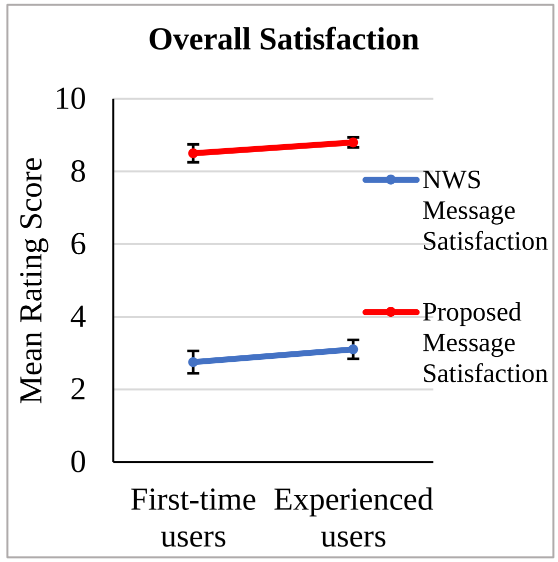
Figures 11 show graphical comparisons between existing and proposed messages with respect to each user group mean rating on each survey item. Overall, the results revealed that users substantially preferred the proposed STW message to the existing NWS STW message, with higher mean rating on the proposed message compared to that on the existing message.



(a) (b)



(c)



(d)

Figure 11. First-time and experienced users' mean rating for each STW survey item: (a) header information, (b) use of delimiter, (c) letter format, and (d) overall satisfaction

A Two-Way Mixed Design ANOVA test was performed to examine the effect of each survey item (header information, use of delimiter, letter format, and overall satisfaction) as a within-subject factor, user group as a between-subject factor, and the interaction between survey items and user group. As shown in Table 3, there was a significant main effect among the levels of each survey item ($p < .05$), in terms of the mean rating score. On the other hand, there was no significant main effect for user group factor ($p > .05$), meaning that both user groups performed similarly. In addition, there was no significant interaction effect between each survey item and user group, indicating that the users' ratings do not depend on their experience levels.

Table 3. Mixed design ANOVA output for the STW weather alert messages

	Mixed Design ANOVA Output				
	Sum of Squares	*Mean Square*	*F*	*DF*	*P*
Header Information	924.80	924.80	884.09		< .001
User Group	1.25	1.25	1.34	1	.255
Header * Group	.45	.45	.43		.516
Use of Delimiter	638.45	638.45	363.46		< .001
User Group	.20	.20	.08	1	.774
Delimiter * Group	1.80	1.80	1.03		.318
Letter Format	638.45	638.45	354.95		< .001
User Group	6.05	6.05	3.78	1	.059
Letter Format * Group	.20	.20	.11		.741
Overall Satisfaction	655.51	655.51	579.63		< .001
User Group	2.11	2.11	1.65	1	.207
Satisfaction * Group	.01	.01	.01		.917

- **Survey Comparison of Weather Alert Messages: WA alert messages**

Figures 12 show graphical comparisons between existing and proposed messages with respect to each user group mean rating on each survey item. Similar to the findings on the STW alert messages, the results revealed that users substantially preferred the proposed WA message to the existing NWS WA message, with higher mean rating on the proposed message compared to that on the existing message.

A Two-Way Mixed Design ANOVA test was performed to examine the effect of each survey item (information location, word expression, use of terminology, and overall satisfaction) as a within-subject factor, user group as a between-subject factor, and the interaction between survey items and user group. As shown in Table 4 and similar to the STW results, there was a significant main effect among the levels of each survey item ($p < .05$) on the WA message, in terms of the mean rating score. On the

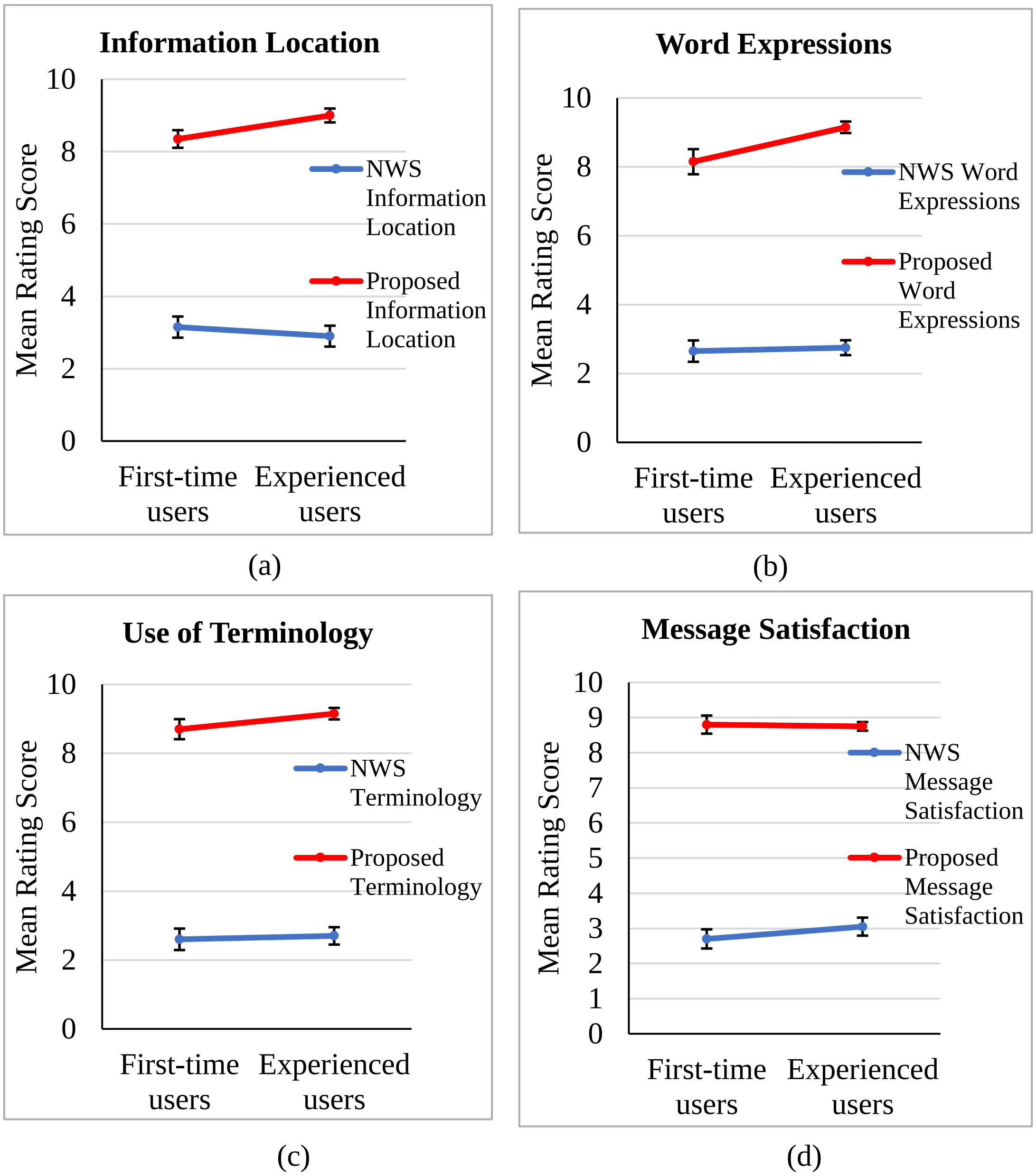


Figure 12. First-time and experienced users' mean rating for each WA survey item: (a) information location, (b) word expression, (c) use of terminology, and (d) overall satisfaction

other hand, there was no significant main effect for user group factor ($p > .05$), meaning that both user groups performed similarly. In addition, there was no significant interaction effect between each survey item and user group, indicating that the users' ratings do not depend on their experience levels.

Table 4. Mixed design ANOVA output for the WA weather alert messages

	Mixed Design ANOVA Output				
	Sum of Squares	*Mean Square*	*F*	*DF*	*P*
Information Location	638.45	638.45	421.93	1	< .001
User Group	.80	.80	.70		.41
Location * Group	4.05	4.05	2.68		.11
Word Expression	762.61	762.61	569.62	1	< .001
User Group	2.11	2.11	2.18		.148
Expression * Group	.101	1.01	.76		.390
Use of Terminology	787.51	787.51	530.83	1	< .001
User Group	1.51	1.51	1.21		.278
Terminology * Group	.61	.61	.41		.524
Overall Satisfaction	696.20	696.20	678.35	1	< .001
User Group	.45	.45	.38		.542
Satisfaction * Group	.80	.80	.78		.383

- **Exit Survey Results**
    - Most Difficult Tasks

For the most difficult task, 20 participants answered the map settings task, 11 participants answered the alert settings task, and 9 participants answered the location search with the pin on map approach. No participant reported any difficulty when interacting with the locations search (app's text bar).

- Overall Usability of the Weather Radio App

The results showed that the majority of participants rated the overall usability of the Weather Radio app between "Fair" and "Good", (M = 3.08, SD = .76). More specifically, 42.5% of the participants (17 participants) rated the overall usability as "Fair" and 32.5 % of them (13 participants) found the Weather Radio app as "Good". In addition, 25 % of the participants (10 participants) rated the usability of the app as "Poor", while no extreme ratings were reported.

- Comments on the Usability of the Weather Radio App

The responses to this question were analyzed using the qualitative content analysis technique (see Table 5). They were categorized into three major categories: settings, location search, and general. Settings were divided into three sub-categories: map settings, NWS alerts, and general setting comments. The map settings seemed to be problematic to many participants (22 participants) who reported that they were confused about how to get to the sub-menu leading to the map setting options, as there was no indication that the map needed to be tapped in order to be able to see the sub-menu. A few suggestions to this issue were made, such as placing the map settings in the general settings menu after the gear icon is tapped.

Six participants explained that the settings should be modified for better layout and organization. For example, one user suggested placing the settings icon at the top right corner instead of its current location at the bottom right corner.

Nine comments related to NWS alerts were reported. For example, one user reported that it was difficult to enable/disable alerts and sub-alerts as this required a prior step of tapping the "NWS Alerts" icon, and that they did not know the meaning of "NWS".

The analysis also revealed that twenty comments were included as issues in the location search task. Fourteen of the comments were about the difficulty of locating a specific place on the map of the small smart phone display. Six comments were

Table 5. Content analysis of usability comments from both first-time and experienced users

Category		Problem/ Expectation	No. of Comments	Representative Examples
Settings	Map settings	Confusion of getting to the sub-menu	22	1) "It should be easier to get to the map settings." 2) "The map settings should be in the settings tab; I was confused for a while trying to find out where the map settings were."
	NWS alerts	Difficulty finding alerts features	9	1) "Overall, the app was good to use except few options to access such as enable/disable NWS alerts as I couldn't figure out what NWS stands for"
	General	Poor Layout and Organization	6	1) "Settings should be right top corner" 2) "Most of the options are under the settings buttons. Instead of that few buttons/options can be made available on the home screen itself."
Location Search		Confusion of locating a place on the map	14	1) "Map view-ability was so clustery (messy) appearance, it can be improved."
		Frustration of controlling the pin on the map	6	1) "Long press on map should drop a pin/move current pin"
General		Difficulty locating desired features	8	1) "It is not that easy to interact with the app. It has a lot of features, but they seem masked and not easy to understand/find on the app." 2) "Consumes more time to search for options." 3) "It should use material design guideline (google)"

reported about the frustration of controlling the pin on the map. For example, one user suggested that tapping the desired location on the map should automatically move the pin instead of the current requirement of long pressing and holding of the pin until it lifts and then moves to the desired location.

Finally, eight comments were made on the general usability of the Weather Radio app. These comments were concerned with the difficulty of finding desired features. Due to the complex menus and non-intuitive terms, some users found the app difficult to use and time consuming. Furthermore, one of the users suggested considering the material design guidelines created by Google for a better design. The user believed that those guidelines could enhance the usability of the application as they provide simple and intuitive designs.

3.2.4 Discussion of Study I

Even though all users successfully accomplished the given search tasks, the other metrics of the experiment revealed multiple usability problems. Those problems could also be found in several other weather alert apps and in apps with similar inherent features. The issues, implications, and the proposed solutions are discussed below.

For the location search task, both first-time and experienced users were significantly slower in using the pin feature on the map than in typing the address within the text bar. This was possibly due to the multiple steps that were required to use the pin feature (see section 3.1 for details) and the need to frequently search the map for a desired location. In addition, a counterintuitive step (i.e. no explanation that the pin on the map had to be pressed for more than one second to move it) further slowed the task completion time.

Using the pin feature in computer display with a mouse may be beneficial as it is easy and intuitive to click on, hold, and drag the pin to the desired location. Specifically, for a very limited number of times, users need to zoom in and out to find the location of interest on larger computer displays compared to that on small smart phone displays. This implies that using the pin feature may not be the most efficient option when searching for a location on a smart phone display. However, even though this study revealed that typing the address within the text bar was much faster than the using the pin on the map for both user groups, users may not always know the exact address of the desired location. Typing a familiar location, such as one might in the Google Map's app, with effective auto suggestions might further enhance the efficiency of the location search feature.

The alert settings task was problematic to first-time users compared to experienced users for two reasons. First, based on their responses to the exit survey as well as their performance during the direct observation, they were confused about which option to choose to find the alert settings menu. Most users kept randomly clicking on each of the available setting options (see Figure 6 (b)) since they could not figure out the meaning of "NWS". Second, the large available number of alerts and sub-alerts within the NWS alert options slowed the participants' performance as they spent much time navigating through some alert menus (see Figure 6 (c) & (d)). It might have helped the users if there was a filtering option that only showed the most critical and widely used alerts and sub alerts, as well as avoiding jargon and unclear abbreviations to enhance the user's experience. These recommendations are also believed to be useful for non-weather apps as understandability of displayed information (Panach et al., 2008) and inclusion of the least amount of menu options required for accomplishing tasks (Whitenton, 2016) are among the top usability requirements.

The map settings task was extremely challenging for first-time users. This was obvious as the first-time users needed a substantially longer time to complete this task compared to experienced users. In addition, the qualitative content analysis showed that most of the users' concerns and comments regarding the usability of the app's features were on the map settings task. The issues with this task were attributed to the included counterintuitive steps. Specifically, users were required to tap the information icon, labeled "i" in a secondary hidden menu, that would appear on the screen if the tiny map portion (see Figure 7 (b)) was tapped. Based on the first-time users' responses to the exit survey and the findings from the direct observation, they struggled a lot with this feature as there was no explanation on how to reach the secondary menu. In addition, finding the map setting options through the information icon "i" was completely unexpected as this icon is commonly used for showing some information about an entire app. Hence, it is recommended to enhance the visibility of all features and clearly indicate their functionality, with particular focus to the map feature on the Weather Radio app. In addition, the app's developers should consider creating a better representative icon of the map settings menu and keeping the "i" icon for displaying information about the app. Creating highly intuitive interfaces would lessen first time users' confusion and greatly enhance the overall usability.

The proposed versions of the weather alert messages (the STW message and the WA message) yielded significantly higher rating scores than the original messages by both experienced and first-time users because the users clearly stated the lack of clarity and organization of the original messages. For example, the severe thunderstorm watch message included several undefined codes in the header information, such as “OKC015, TXC009...” Users could not understand and probably did not need to know that those were the geographical codes of the names of the areas under alert. Another example is the description of the wind impact in the wind advisory message, which uses technical information, such as “SOUTH TO SOUTHWEST 25 TO 35 MPH WITH GUSTS 40 TO 50 MPH.” Such technical representation of information was not understandable based on their low mean rating score shown in the result section. It is worth mentioning that the alert messages received by end users follow the process shown in Figure 13. Specifically, once NWS devices detect extreme weather conditions, they issue corresponding alert messages; third parties, including smart phone weather apps directly receive them and push them automatically to end-users.

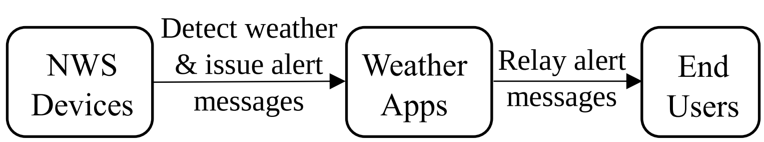


Figure 13. Current method of delivering alert messages to end users

Hence, it would be beneficial if NWS considered sending user friendly alert messages or providing a guideline that would allow weather app developers to modify the original alert messages so that they facilitate easier comprehension. Failing to fully comprehend warning messages or alerts of any time-critical system, such as a smart phone weather system, in a timely manner may significantly impact users’ lives.

3.3 Study II: Eye Tracking Analysis of Smart Phone Weather Applications

3.3.1 Objective

In this study, we aim to examine whether the eye tracking can better support the analysis from the findings of Study I. Specifically, the goal of this study is to objectively understand users' cognitive processes and decision-making strategies, based on their eye movements, when performing tasks on smart phone weather apps. Similar to Study I, in this study we collected data from both first-time and experienced users on the same smart phone weather app, Weather Radio. As alert messages cannot be accessed on weather apps unless in case of active weather alert conditions, we did not consider the alert messages task in this experiment. The three features tested in this study are location search, alert settings, and map settings. The findings from this study are believed to provide a solid foundation regarding the use of eye tracking in the area of usability evaluation of apps on smart phone devices with particular significance in the weather domain.

3.3.2 Method

Participants

Forty undergraduate and graduate students (males and females) were recruited at the University of Oklahoma (Norman campus). All participants owned smartphones and were regular smartphone (iPhone) users prior to the experiment. The age of participants varied from 21 to 44 with a mean of 24.7 (*SD* = 4.9). All participants had no prior experience with using the Weather Radio app (version: 3.0.5). In the experiment, the participants were randomly assigned into two groups. The first group [20 trained (experienced) participants] received complete training on using all the app's functions by one of the study researchers, while the second group [20 untrained (first-time) participants] received no training at all. The data of the experienced participants was considered as a reference measure of comparison with the first-time participants' data.

Materials

Hardware specification (a smart phone "iPhone 6", a Tobii Pro TX300 eye tracker with four built-in cameras, and a 19-inch Liquid-Crystal-Display "LCD" monitor) and software specification (Tobii studio software "version 3.3" and a Personal Computer "PC" emulator receiver called "Reflector 2") were used in this experiment. Specifically, as the eye tracker was attached to the bottom of the monitor in front of the user, we placed the smart phone, by taping it, on the monitor. The Weather Radio app was emulated and displayed right behind the smartphone on the monitor using Reflector 2; the emulated app display had exactly the same dimensions as the smartphone display with no angle errors (see Figure 14). All users' physical interactions with the Weather Radio app's features on the smart phone device were mirrored on the emulated display in real time and users' movements appeared on the emulated Weather Radio interface. The accuracy of the eye tracker was 0.5° of visual angle and the data was collected at 300 Hz. The Tobii studio was used to obtain the users' eye tracking data.

Figure 14. Experimental setup showing an example of how a user's eye fixation (yellow circle) on the actual smartphone appear on the emulated interface with the same distance (d) from a reference point

Procedure and Tasks

The experiment took place in the University of Oklahoma (Norman campus) in Carson Engineering Center, room 23. The participants were asked to sign an informed consent form before beginning the experiment, as well as complete a pre-experiment demographic questionnaire. The experienced group was introduced to the Weather Radio app and given unlimited time to familiarize themselves with the app's features. Once a participant felt ready to begin the experiment, the participant would verbally state "I am ready to begin the experiment". On average, the experienced group took 8.43 minutes to complete the training session. Prior to beginning the experiment, a simple calibration of participants' eye movements by observing a red moving dot on the screen was performed. All participants were then asked to begin the experiment by performing three specified tasks: location search, alert settings, and map settings. The participants were asked to accomplish the given tasks as quickly as possible. The experiment's tasks were counterbalanced between participants.

The tasks used in Study II were exactly the same as the search tasks used in Study I. For the location search, two approaches towards the location search were considered in this study: 1) controlling the pin on the map and 2) typing the location address in the app's text bar. Specifically, the participants were asked to find the location of "Mount Auburn Hospital in Cambridge, MA" using the two approaches for comparison purposes. For the alert settings, the participants were asked to enable (turn on) *Tornado Warning* and *Severe Thunderstorm Warning* and disable (turn off) *Tornado Watch* and *Severe Thunderstorm watch*. For the map settings, the participants were asked to change the Map Type from *Standard* to *Hybrid* and Weather Layer from *Radar* to *Cloud*. See section 3.1 for details.

Variables

Two independent variables were included in this study and each included two levels. The first independent variable (approach type) was used for only the location search task and included two levels: pin on map and app's text bar. The second independent variable (experience) was used for all three tasks and included two levels: experienced group and first-time group. The dependent variables for all the given tasks were the same: duration of eye fixations, number of eye fixations, and scanpath patterns.

Data Analysis

The experiment's data were analyzed with respect to pre-determined AOIs. Specifically, the Weather Radio app's interface was divided into six areas of interests (AOIs): 1) Overall, which represents the whole display during the users' entire interaction; 2) Header (H), which shows the top part of the displayed screen; 3) Text Bar (TB), which shows where to type and search for location; 4) Main Display (MD), which shows the main part of the displayed screen; 5) Soft Keypad (SK), which shows the keypad used to input letters and numbers; 6) Bottom Menu (BM), which shows the menu used to access the app's features (see Figure 15).

Parametric tests were performed to analyze the differences among the levels of the location search task with respect to the overall and specific AOIs. For the alert settings and map settings tasks, non-parametric tests were used to determine the differences between first-time and experienced users for the overall and specific AOIs; the non-parametric tests were used, as the alert settings and map settings tasks' data did not meet the assumptions of the parametric tests. In addition, a transition matrix analysis represented by AOI weighted transition diagrams was used to illustrate the users' eye transition activities among the specific AOIs. Finally, correlation tests were used to determine the association between the eye tracking metrics' data for each task.

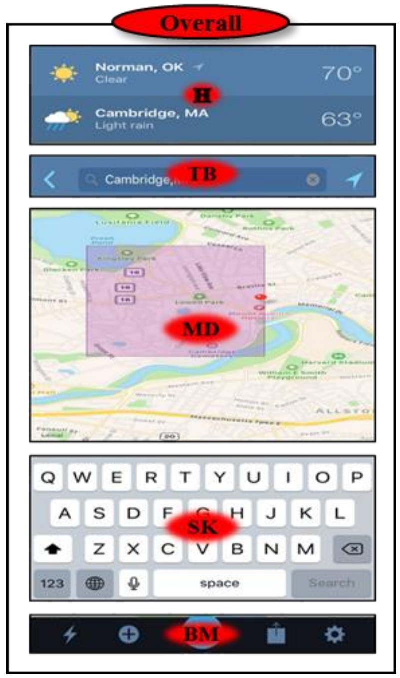
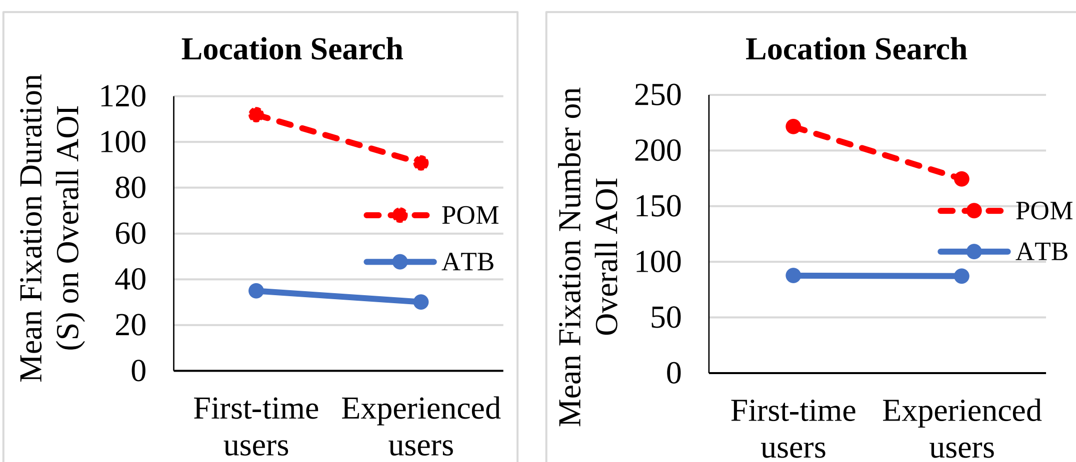


Figure 15. Example of overall and specific AOIs of the Weather Radio app

3.3.3 Results

- **Weather Radio app tasks results (overall AOIs)**

  - **Location Search**



(a) (b)

Figure 16. Differences between first-time and experienced users on the location search task in terms of the overall eye fixations duration (a) and number (b)

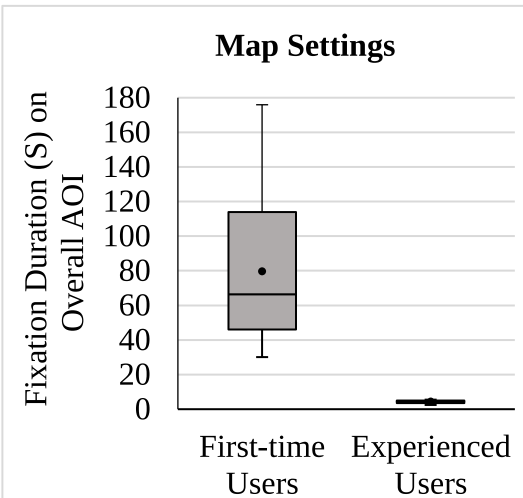
Figure 16 shows a graphical comparison between first-time and experienced users on the location search task, in terms of both eye fixations duration and number.

A Two-Way Mixed Design AVOVA with approach type (pin on map and app's text bar) as a within-subjects factor and user group (first-time and experienced users) as a between-subjects factor was conducted in terms of the overall fixation durations. The results showed that there was a significant main effect for approach type (F (1, 38) = 70.87, $p < .001$) on the overall fixation durations, with significantly shorter overall fixation durations on the app's text bar (M = 32.51s) compared to that on the pin on map (M = 101.36s). In contrast, there was no significant main effect for user group (F (1, 38) = 2.29, $p = .138$) on the overall fixation durations, with first-time (M = 73.43s) and experienced users (M = 60.45s) performing similarly. In addition, there was no significant interaction effect between approach type and user group (F (1, 38) = .99, p = .326), meaning that the overall eye fixation durations observed on each location search approach does not depend on whether the users are experienced or not.

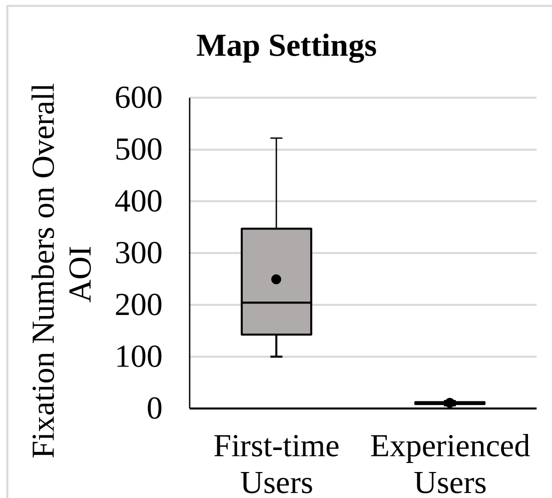
Similarly, the Mixed Design ANOVA with the same factors was conducted, in terms of the overall fixation numbers. It was revealed that there was a significant main effect for approach type (F (1, 38) = 71.85, $p < .001$) on the overall fixation numbers, with significantly less overall fixation numbers on the app's text bar (M = 83.10) compared to that on the pin on map (M = 197.95). In contrast, there was no significant main effect for user group (F (1, 38) = 3.70, $p = .062$) on the overall fixation numbers, with users showing similar average fixations numbers for first-time (M = 154.48) and experienced users (M = 126.58). In addition, there was no significant interaction effect between approach type and user group (F (1, 38) = 1.97, p = .169), meaning that the overall eye fixation numbers observed on each location search approach does not depend on whether the users are experienced or not.

- **Alert Settings**

Figure 17 shows a graphical comparison between first-time and experienced users on the alert settings task, in terms of both eye fixations duration and number.



(a)



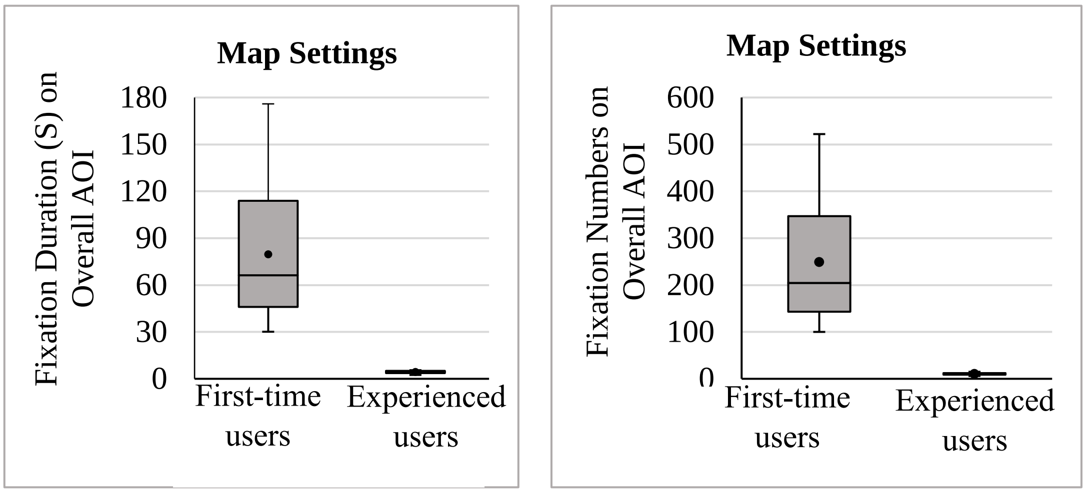
(b)

Figure 17. Differences between first-time and experienced users on the alert settings task in terms of the overall eye fixations duration (a) and number (b)

A Mann-Whitney U test was conducted to examine the difference between first-time and experienced users, in terms of the median overall eye fixation durations and numbers on the alert settings task. For the overall eye fixation durations, the results illustrated that there was a significant difference between the two user groups (U = 17, p < 0.001), where the experienced users spent significantly shorter overall eye fixation durations than the first-time users. Similar to the eye fixation durations, there was a significant difference ($U = 21$, $p < 0.001$), with the experienced users making significantly less overall eye fixation numbers than the first-time users.

- **Map Settings**

Figure 18 shows a graphical comparison between first-time and experienced users on the map settings task, in terms of both eye fixations duration and number.



(a) (b)

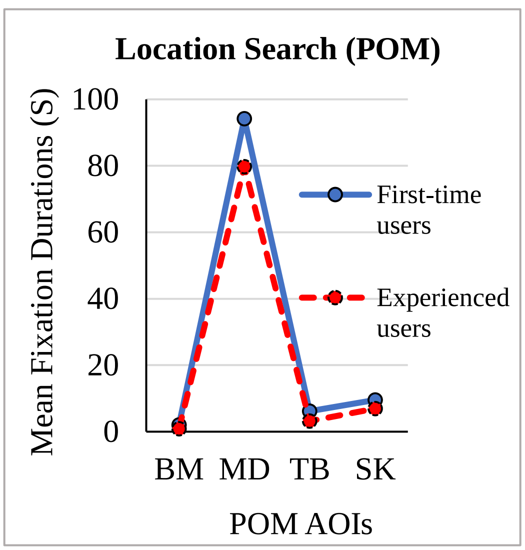
Figure 18. Differences between first-time and experienced users on the map settings task in terms of the overall eye fixations duration (a) and number (b)

The Mann-Whitney U test revealed a statistically significant difference between the two groups ($U = < 0.001$, $p < 0.001$), where the experienced users spent significantly shorter overall eye fixation durations than the first-time users on the map settings task. Similarly, there was a significant difference ($U = < 0.001$, $p < 0.001$), with significantly less overall eye fixation numbers by the experienced users.

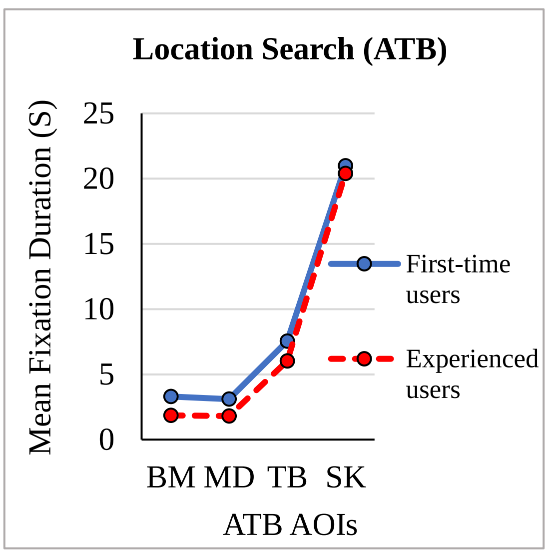
- **Weather Radio app tasks results (specific AOIs)**
  - **Location Search**

Figure 19 shows a graphical comparison between first-time and experienced users on the location search task, in terms of mean eye fixations duration and number; (a & c) in terms of the pin on map (POM) and (b & d) are in terms of the app's text bar (ATB).

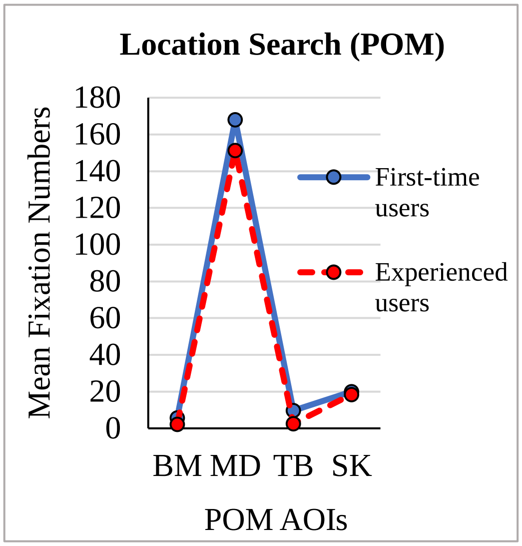
Mixed Design ANOVA tests were performed to examine the effects of the specific used AOIs (BM, MD, TB, and SK) on both location search types as a within-subject factor, user groups as a between-subject factor, and the interaction between AOIs and user groups. In addition, post-hoc analyses using Tukey's tests were conducted to determine the differences among the used AOIs. The data were investigated in terms of both the eye fixation durations and numbers.



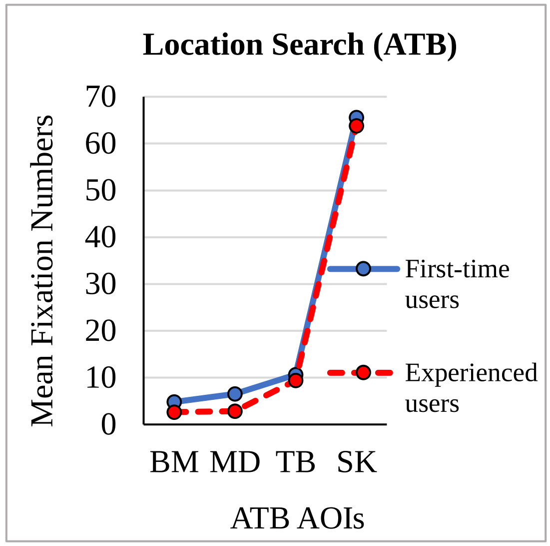
(a)



(b)



(c)



(d)

Figure 19. Differences between first-time and experienced users on location search approaches: POM (a & c) and ATB (b & d), in terms of the mean eye fixations duration (a & b) and number (c & d)

Pin on map (POM)

The results showed that there was a significant difference across the four AOIs, F (3, 114) = 102.83, p <.001, while no significant difference was observed between the user groups, F (1, 38) = 1.62, p = .211, in fixation durations. There was also no significant interaction between POM AOIs and user groups, F (3, 114) = .57, p = .637. The post hoc analysis revealed that all POM AOIs were significantly different from each other (p < .05).

In addition, it was shown that there was a significant difference across the four POM AOIs, F (3, 114) = 142.75, p <.001, while no significant difference was observed between the user groups F (1, 38) = 2.88, p = .098, in fixation numbers. There was also no significant interaction between POM AOIs and user groups, F (3, 114) = 1.35, p =

.263. The post hoc analysis revealed that all POM AOIs were significantly different from each other ($p < .001$).

App's text bar (ATB)

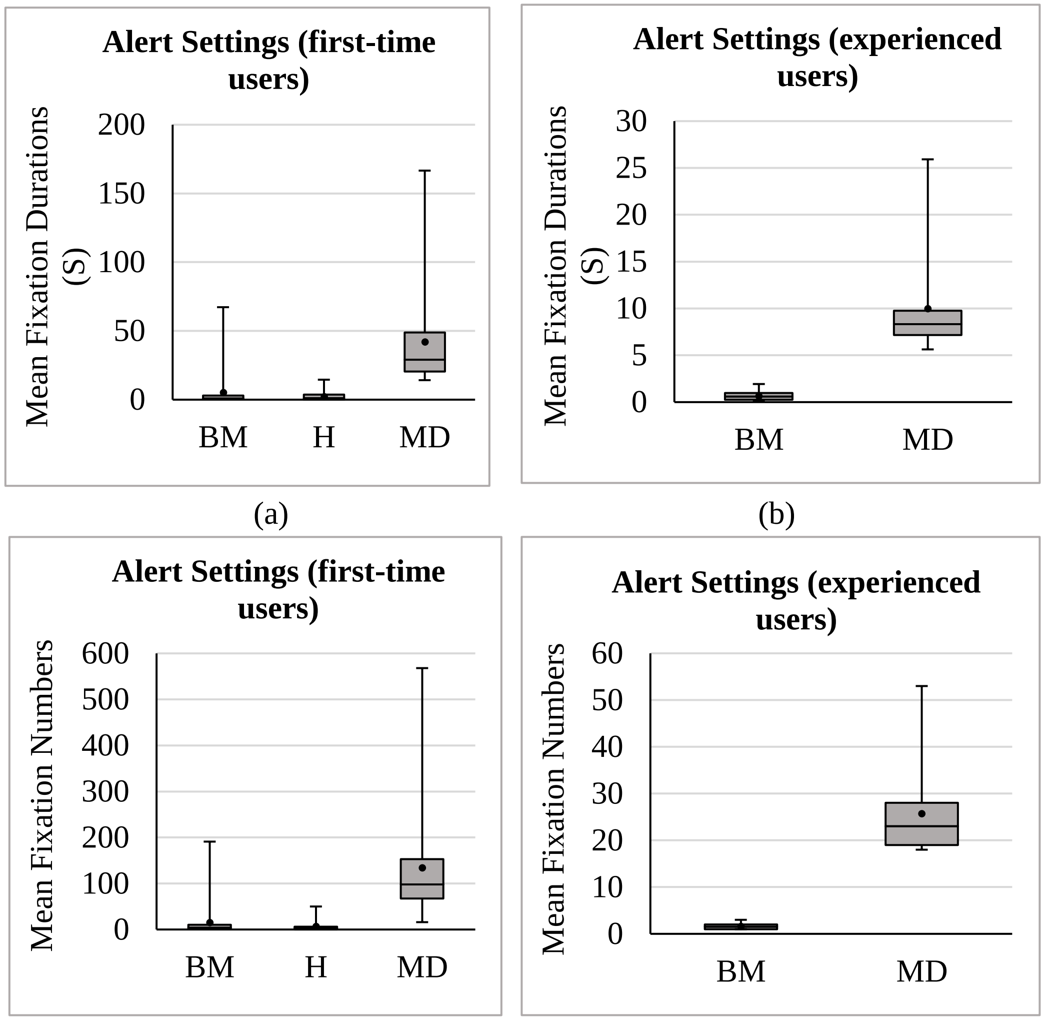
For the ATB, the results showed that there was a significant difference across the four AOIs, F (3, 114) = 200.21, p <.001, while no significant difference was observed between the user groups F (1, 38) = 4.22, p = .057, in fixation durations. There was also no significant interaction between ATB AOIs and user groups, F (3, 114) = .12, p = .949. The post hoc analysis revealed that all ATB AOIs were significantly different from each other (p < .05), except the AOI "BM" from the AOI "MD" (p = 1.00).

In addition, it was shown that there was a significant difference across the four ATB AOIs, F (3, 114) = 676.66, p <.001, while no significant difference was observed between the user groups F (1, 38) = 3.77, p = .060, in fixation numbers. There was also no significant interaction between POM AOIs and user groups, F (3, 114) = .21, p = .89. The post hoc analysis revealed that all ATB AOIs were significantly different from each other (p < .05), except the BM from MD (p = 1.00).

- **Alert Settings**

Figure 20 shows a graphical comparison between first-time and experienced users on the alert settings task, in terms of the mean eye fixation durations and numbers.

Friedman tests were conducted to determine if statistically significant differences existed among three AOIs (H, MD, and BM) used by first-time users to complete the alert settings task. On the other hand, Wilcoxon signed-rank tests were



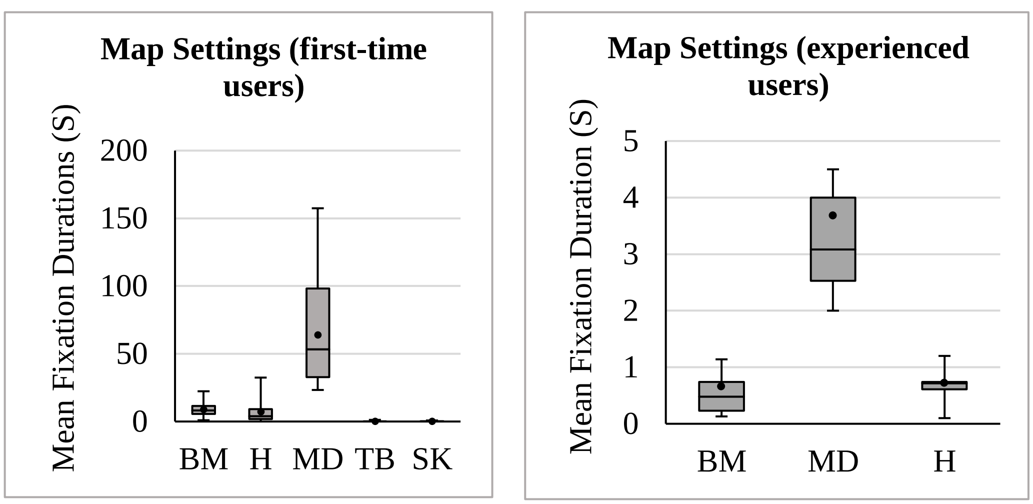
(a) (b) (c) (d)

Figure 20. Differences between first-time and experienced users on the alert settings task, in terms of the mean eye fixations duration (a & b) and number (c & d)

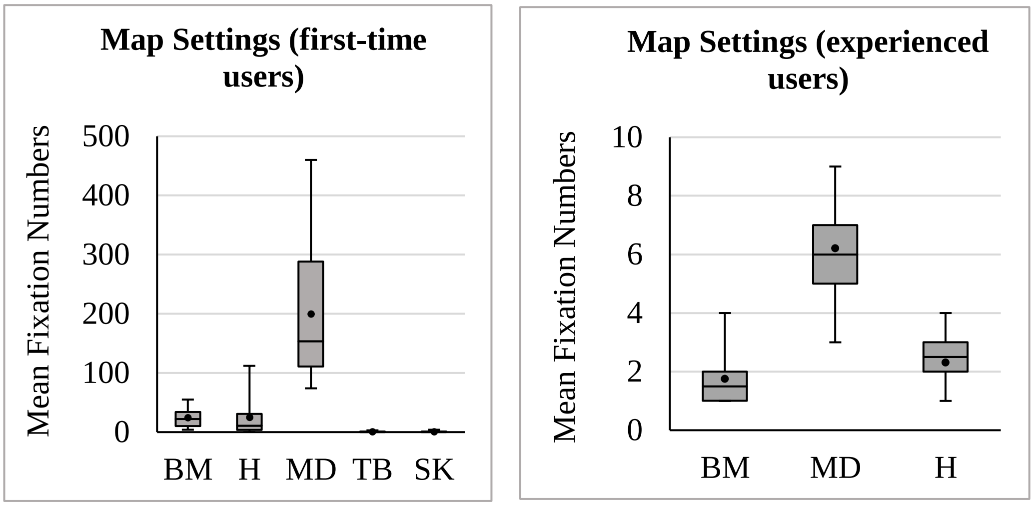
conducted for examining the difference between the AOIs used by experienced users, as they used only two AOIs (BM and MD) to complete this task. Both tests were performed, in terms of the median eye fixations duration and number. For first-time users, the results showed that there were statistically significant differences among the three AOIs in terms of both the eye fixation durations ($\chi 2(2) = 56.73$, p <.001) and eye fixation numbers ($\chi 2(2) = 54.28$, p <.001). To test where the differences exactly occurred, post-hoc analyses of pairwise comparisons with Wilcoxon signed-rank tests were performed with a Bonferroni correction alpha of 0.05/3 = 0.017. It was found that the three AOIs were significantly different from each other, except the AOI “BM” from AOI “H” (Z = -1.248, p = .212) in terms of eye fixation durations and (Z = -1.253, p = .210) in terms of eye fixation numbers, respectively. For the experienced users, the Wilcoxon signed-rank test results revealed that there were statistically significant differences between the AOI “BM” and AOI “MD” in terms of both the eye fixation durations (Z = -3.920, p <.001) and numbers (Z = -3.924, p <.001).

- **Map Settings**

Figure 21 shows a graphical comparison between first-time and experienced users on the map settings task, in terms of the mean eye fixation durations and numbers.



(c) (d)



(c) (d)

Figure 21. Differences between first-time and experienced users on the map settings task, in terms of the mean eye fixations duration (a & b) and number (c & d)

Friedman tests were conducted to determine if statistically significant differences existed among all specific AOIs used by both first-time and experienced users to complete the map settings task. The tests were performed in terms of the median eye fixation durations and numbers. It is worth mentioning that first-time users used all five AOIs, while experienced users used three AOIs: (H, MD, and BM).

For the first-time users, the results showed that there were statistically significant differences among the five AOIs in terms of both the eye fixation durations ($\chi 2(4) = 73.33$, $p < .001$) and numbers ($\chi 2(4) = 73.52$, $p < .001$). The post-hoc analyses with a Bonferroni correction alpha of 0.05/5 = 0.01 revealed that all five AOIs were significantly different from each other, except two pairwise comparisons; the AOI "BM" from AOI "H" and the AOI "TB" from AOI "SK" ($Z = -1.157$, $p = .247$) and ($Z = -.866$, $p = .386$) in terms of eye fixation durations and ($Z = -.805$, $p = .421$) and ($Z = -.212$, $p = .832$) in terms of the eye fixation numbers, respectively. For experienced users, it was revealed that there were statistically significant differences among the used AOIs, in terms of both the eye fixation duration ($\chi 2(2) = 56.13$, $p < .001$) and numbers ($\chi 2(2) = 54.02$, $p < .001$). The post-hoc analyses showed that the AOI "MD" was significantly different from both the AOI "BM" and the AOI "H", in terms of both the eye fixation durations and numbers. However, the results showed that the AOI "H" was similar to the AOI "BM" in terms of the eye fixation durations ($Z = -1.031$, $p = .344$) and numbers ($Z = -1.36$, $p = .128$).

- **Results of scanpath observations**

For more insight regarding users' performances on the Weather Radio app, each user's scanpath on each of the given tasks was analyzed based on their scanpath data (see example of a user's eye movements in Figure 22).

As sequential operational steps are required to successfully complete the experiment's tasks, a diagram (based on the locations of each operational step) of the required scanpath patterns for each task was created and shown under their relevant tasks in the following sections. The diagrams were considered as base standards for comparing each user's scanpath patterns on each of the experiment's tasks, in terms of the users' proportion of performing exactly as the required scanpaths.

In addition, all first-time and experienced users' eye transitions among AOIs (AOI-transition activities) with their average number of occurrences were analyzed using the transition matrix approach and visualized by AOI weighted transition diagrams.

The diagrams' edges (arrows) between the AOIs represent the users' eye

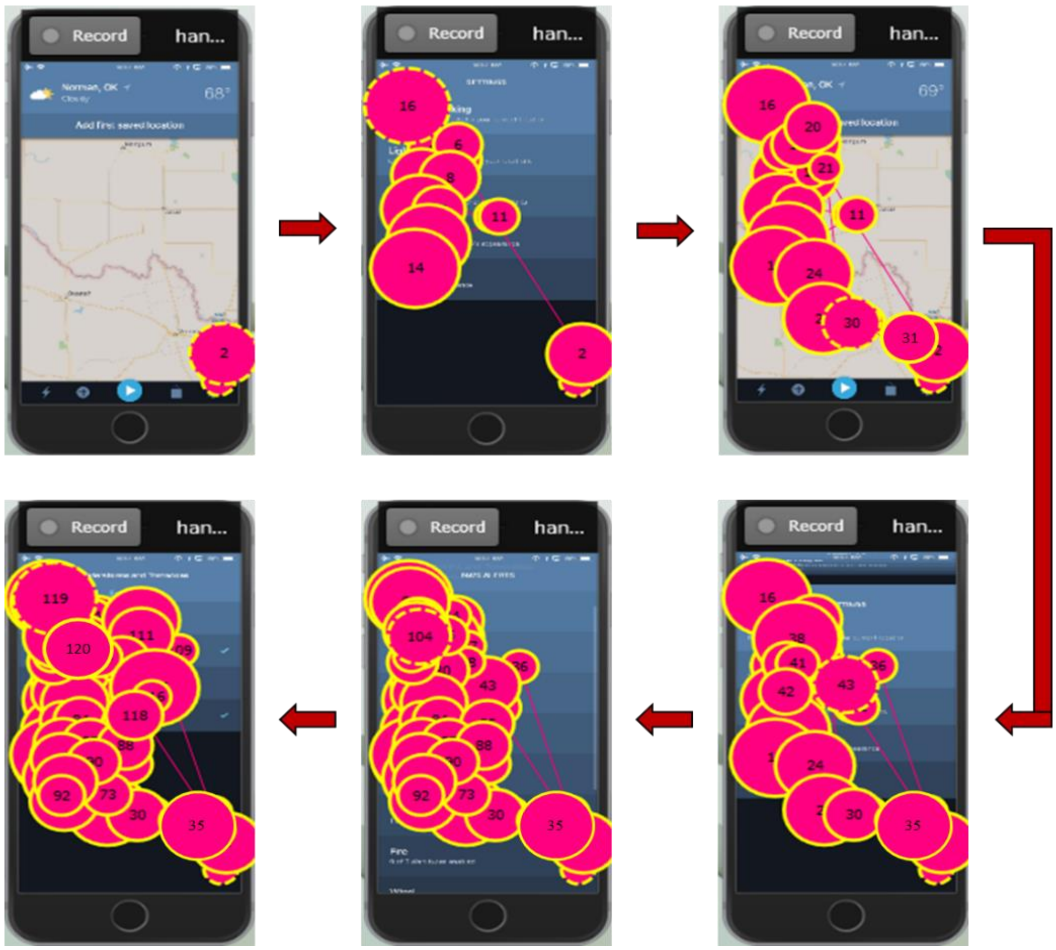


Figure 22. Example of a user's accumulative scanpath on the alert settings task transitions among AOIs and the adjacent numbers are the average numbers of eye transitions. The arrows show the beginning and the end of the transition activities, and the thickness of the arrows is proportional to the average weight; the thicker the arrow the more weight it has.

- **Location Search (pin on map)**

The diagram in figure 23 shows all the possible scanpaths (total of 8 transitions between AOIs) needed to successfully accomplish the location search task using the pin

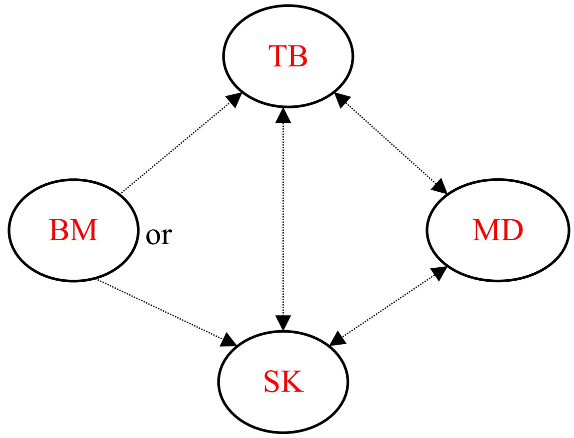
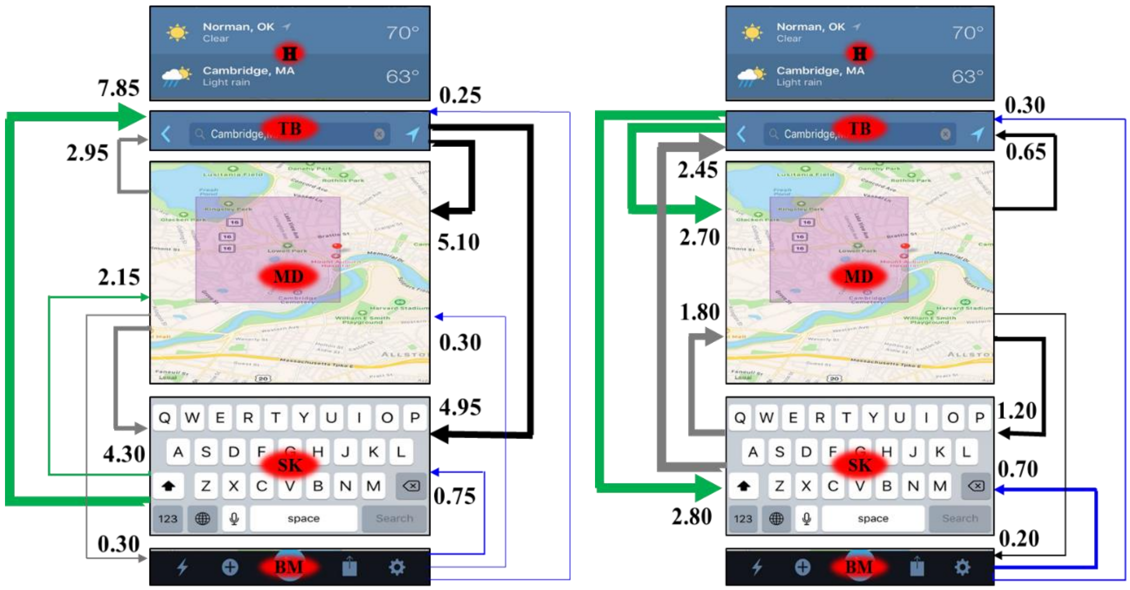


Figure 23. Scanpath pattern needed to successfully complete the location search (pin on map)

on map approach. Specifically, these scanpaths consist of the following: 1) fixating eyes on the AOI “BM” to tap the “+” symbol, 2) shifting eyes to either the AOI “TB” or the AOI “SK”, (depending on the user’s preference) to type desired city & state, 3) shifting eyes back and forth between the AOIs “TB”, “SK”, and “MD” to coordinate typing the city & state and selecting it from the list of options shown in the AOI “MD”, and 4) finally fixating eye on the AOI “MD” to search the map for desired location. For details about the steps required to complete this task, see section 3.1.

Overall, no substantial differences were observed between first-time and experienced users, in terms of the number of users who appeared to follow any of the required scanpaths for this task. Specifically, 80% of experienced users and 70% of first-time users followed the required scanpaths.

The diagrams (a & b) in figure 24 show the actual AOI-transition activities that were made by first-time and experienced users to complete this task. It can be seen that first-time users performed all required transition activities and two additional transitions



(a) (b)

Figure 24. First-time (a) and experienced (b) users' eye transitions among AOIs for location search (pin on map)

to irrelevant AOIs, while experienced users performed the required transition activities and one additional transition activity to an irrelevant AOI. However, those additional transition activities had significantly less average weight compared to the relative transitions.

- **Location Search (app's text bar)**

The diagram in Figure 25 shows all the possible scanpaths (total of 4 transitions between AOIs) required to successfully accomplish the location search task using the app's text bar. These scanpaths consist of the following: 1) fixating eyes on the AOI "BM" to tap the "+" symbol, 2) shifting eyes either to the AOIs "TB" or "SK" in the next screen to begin typing the desired full address, and 3) finally shifting eyes back and forth between the AOIs "TB" and "SK" when typing the address.

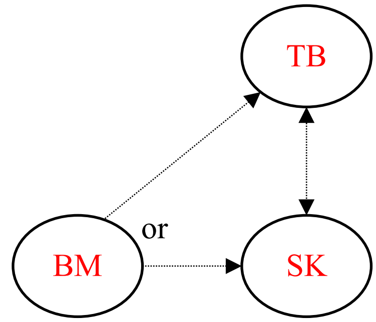
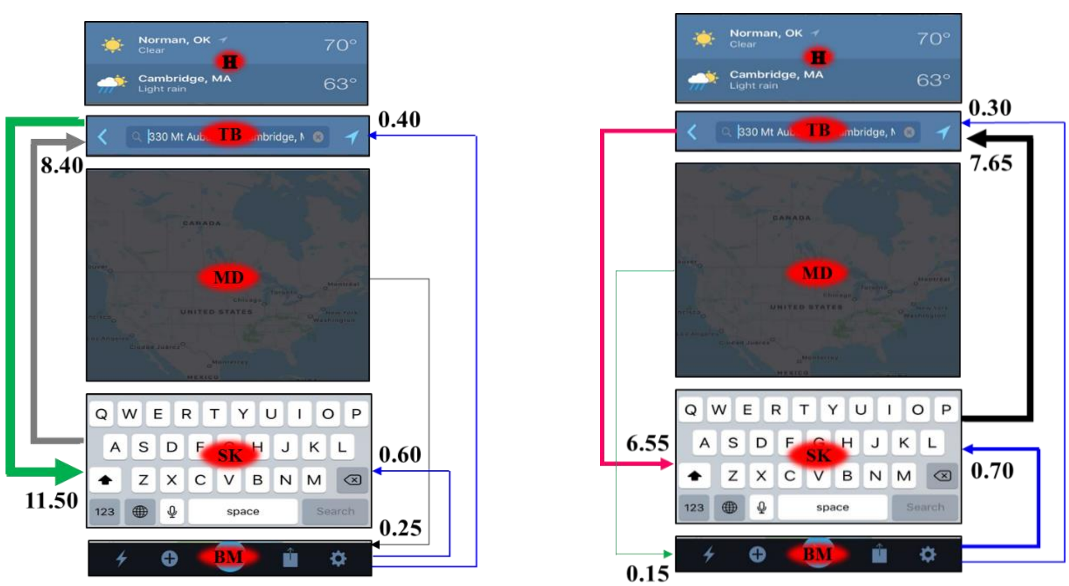


Figure 25. Scanpath pattern needed to successfully complete the location search (app's text bar)

Similar to the location search task (pin on map), even though a larger percentage of experienced users followed the required scanpaths as compared to first-time users on the app's text bar, no major differences were observed. Specifically, 85% of the experienced users and 75% of the first-time users performed the required scanpaths.



(a) (b)

Figure 26. First-time (a) and experienced (b) users' eye transitions among AOIs for location search (app's text bar)

The diagrams (a & b) in Figure 26 illustrate that both first-time and experienced users performed all four transition activities, as well as one additional transition activity to an irrelevant AOI for each user group. However, in comparison to its relative transitions, that additional transition had significantly less average weight.

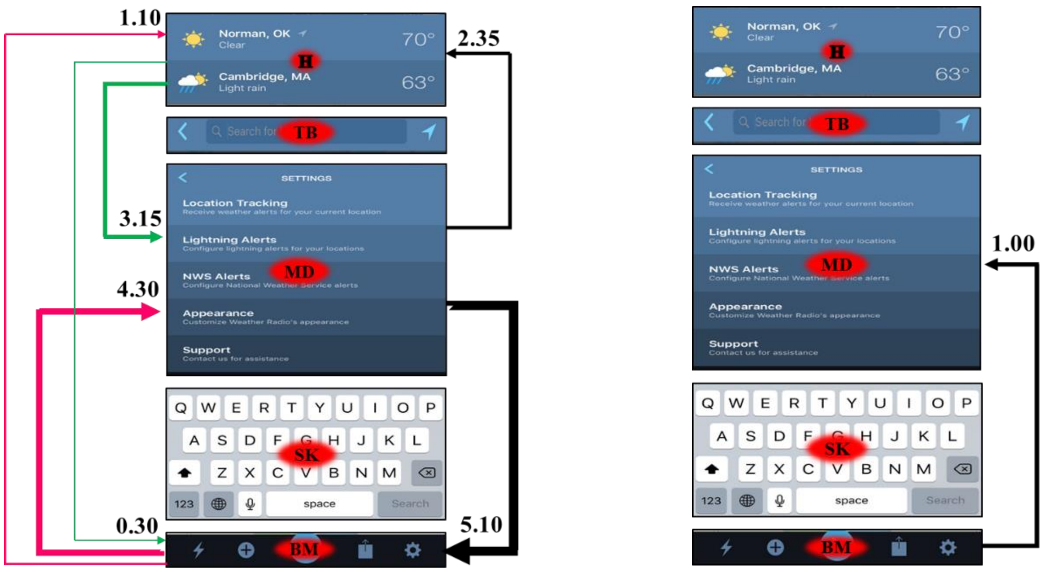
- **Alert Settings**

Figure 27. Scanpath pattern to successfully complete the alert settings task

The diagram in figure 27 shows the scanpath (only one transition activity between AOIs) needed to successfully accomplish the alert settings task. This scanpath is to first fixate eye on the AOI "BM" to tap the gear symbol and then shift eye to the AOI "MD" in the next screen to access the alert settings menu. For details about the steps required to accomplish this task, see section 3.1.

The scanpath results revealed that there was a substantial difference between the two user groups, in terms of the number of users who exactly followed the required scanpath for this task. Specifically, all experienced users (100%) exactly followed the required scanpath, while only 20% of the first-time users followed the required scanpath.

The diagrams (a & b) in Figure 28 show the actual AOI-transition activities that were made by first-time and experienced users to complete this task. The diagrams show that the experienced users performed the only required transition, while the first-time users performed that required transition activity, as well as 5 additional unnecessary transition activities among AOIs with different average weights to



(a) (b)

Figure 28. First-time (a) and experienced (b) users' eye transitions among AOIs for alert settings task

complete this task. Figure 25 (a) also shows that the first-time users made numerous back and forth eye transitions, especially between the AOIs "BM" and "MD".

- **Map Settings**

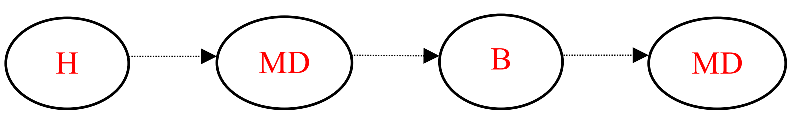


Figure 29. Scanpath pattern to successfully complete the map settings task

The diagram in Figure 29 shows the scanpath (total of three transition activities between AOIs) needed to successfully accomplish the map settings task. This scanpath is to 1) fixate eyes on the AOI "H" to tap the desired saved location for which a user needs to access the map and control its settings, 2) shift eyes to the AOI "MD" to tap the map portion shown in daily forecasts, 3) shift eyes to the AOI "BM" to tap the "i" symbol, and 4) finally shift eyes back to the AOI "MD" to access the map settings menu. For details about the steps required to accomplish this task, see section 3.1.

For the scanpath results, extremely different scanpath patterns between the two user groups were observed. Specifically, 100% of the experienced users completely followed the required scanpath for this task, while none of the first-time users followed the required scanpath.

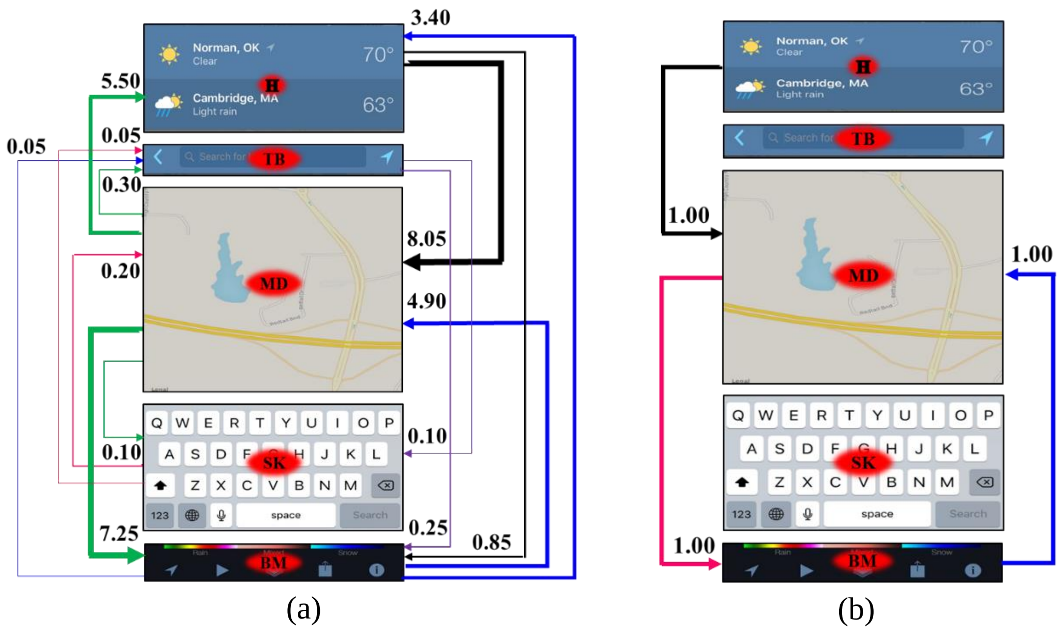


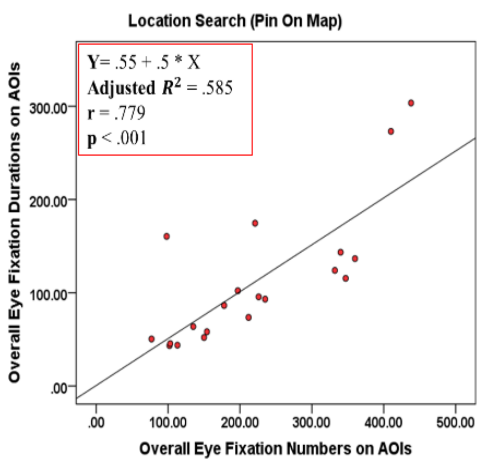
Figure 30. First-time (a) and experienced (b) users' eye transitions among AOIs for map settings task

The diagrams (a & b) in figure 30 show the actual AOI-transition activities that were made by first-time and experienced users to complete this task. The diagrams illustrate that experienced users performed only the required transitions, while first-time users performed the required transition activities, as well as 10 additional unnecessary transition activities to complete this task. The first-time users' transition diagram clearly shows that they were randomly transitioning among AOIs while searching for needed information.

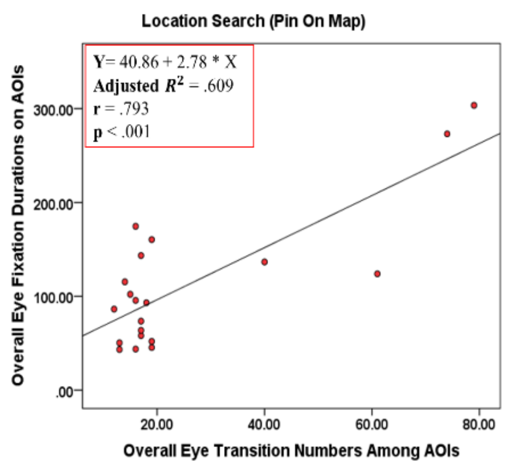
- **Results of pairwise correlations between the eye tracking metrics**

Parametric linear regression with Pearson's correlation tests (for data of location search task) and non-parametric linear regression with Spearman's tests (for data of alert settings and map settings tasks) were performed to assess the relationship between eye fixation numbers on AOIs, eye fixation durations on AOIs, and eye transition numbers among AOIs for each user group. The correlation coefficients were categorized into two qualitative strength measures, strong or weak. Specifically, a correlation between two variables is considered to be strong when the correlation coefficient (r) is 0.5 or above. On the other hand, a weak correlation is determined when the correlation between two variables yields a correlation coefficient of 0.49 or less.

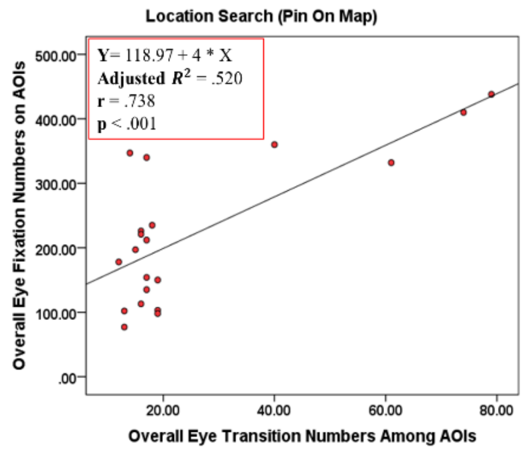
- **Location Search (pin on map)**



(c)



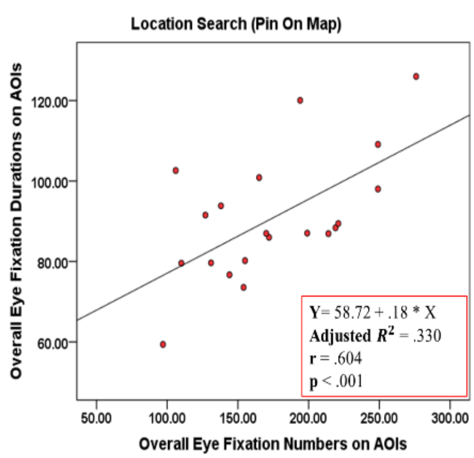
(d)



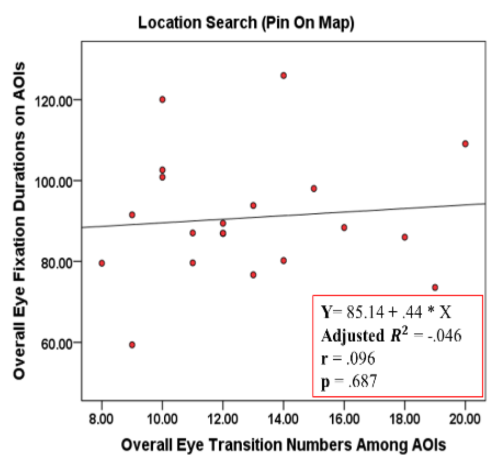
(c)

Figure 31. Scatterplots of the correlations between (a) fixation durations and numbers, (b) fixation durations and transition numbers, and (c) fixation numbers and transition numbers for first-time users on the location search task (pin on map)

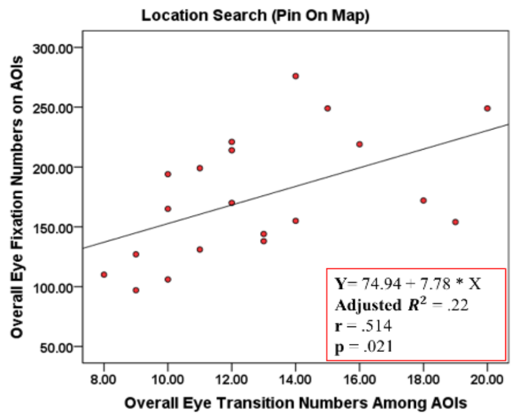
As shown in Figure 31, the location search (pin on map) results for first-time users show that there were strong positive correlations ($r > 0.5$) among all correlation combinations of eye tracking metrics, meaning that an increase in one variable resulted in an increase in the other variable.



(a)



(b)

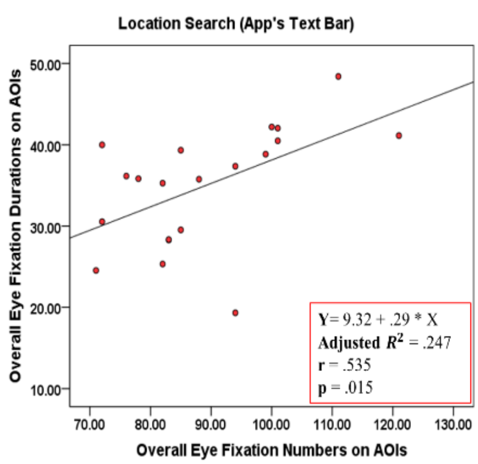


(c)

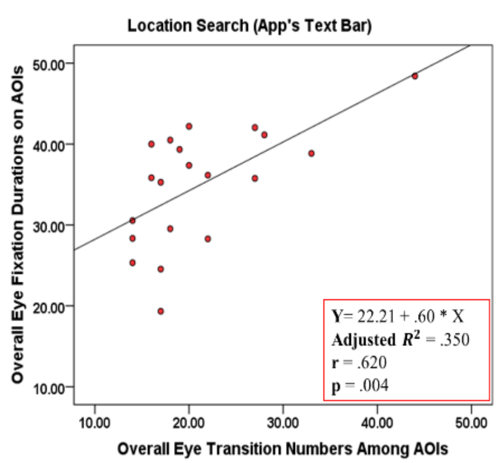
Figure 32. Scatterplots of the correlations between (a) fixation durations and numbers, (b) fixation durations and transition numbers, and (c) fixation numbers and transition numbers for experienced users on the location search task (pin on map)

For experienced users, Figure 32 shows that barely strong correlations were observed between “fixation durations and numbers” and between “fixation number and transition numbers”, while very weak correlation occurred between “fixation durations and transition numbers.”

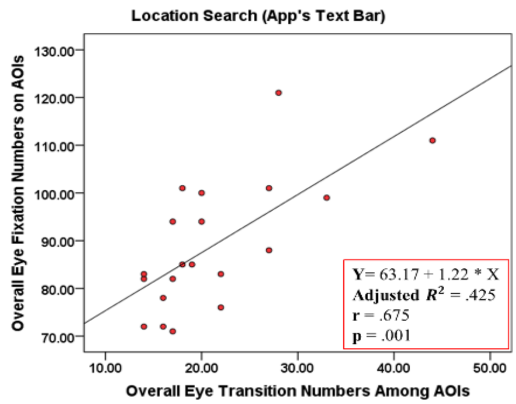
- **Location Search (app’s text bar)**



(a)



(b)

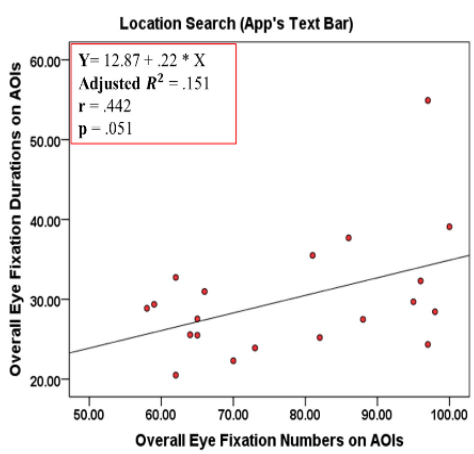


(c)

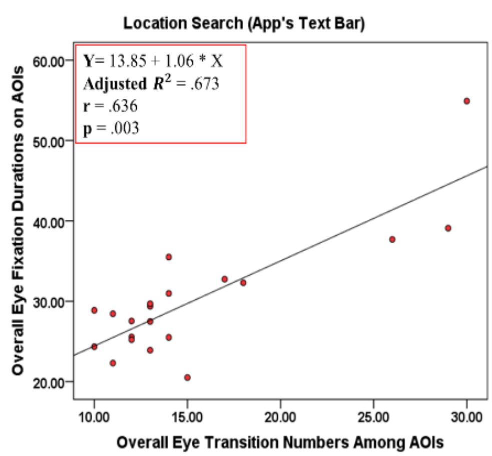
Figure 33. Scatterplots of the correlations between (a) fixation durations and numbers, (b) fixation durations and transition numbers, and (c) fixation numbers and transition numbers for first-time users on the location search task (app's text bar)

As shown in Figure 33, there were strong positive correlations among all correlation combinations of eye tracking variables for first-time users.

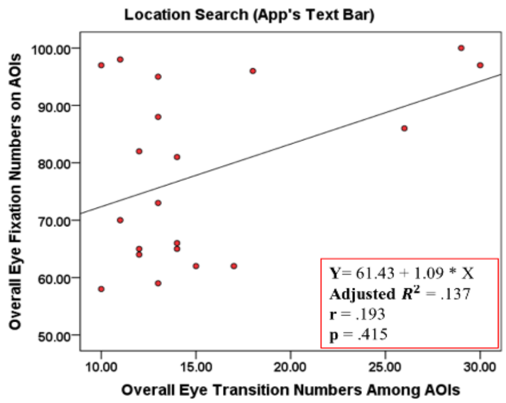
Figure 34 shows that there were weak correlations among all combinations of the eye tracking variables on the location search task (app's text bar) for experienced users, except among "fixation duration and transition numbers."



(e)



(f)

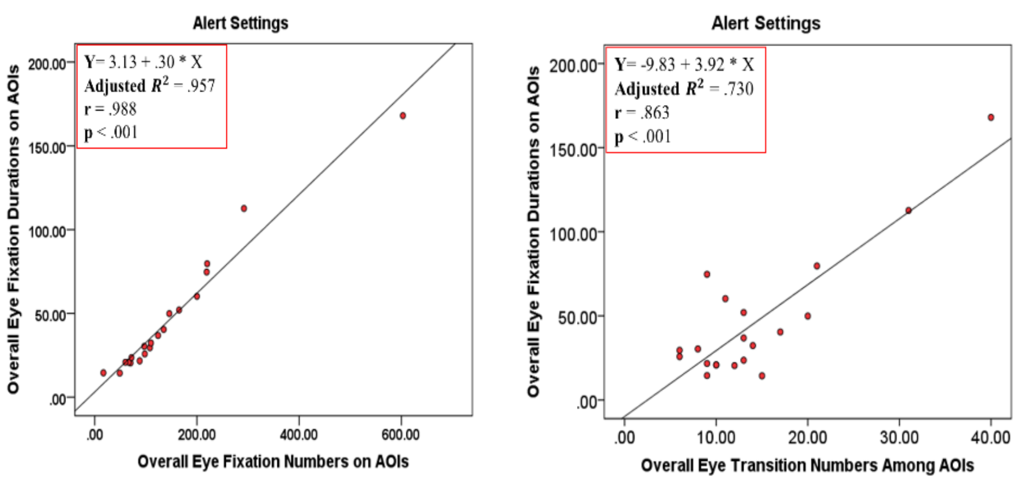


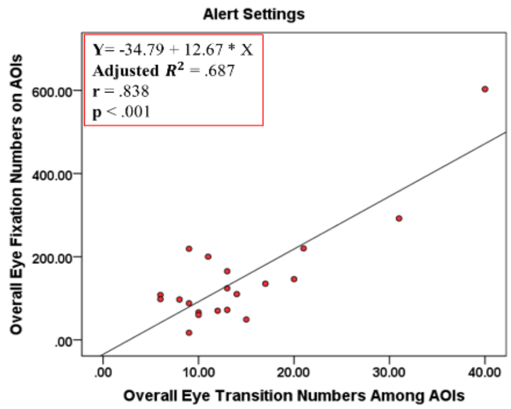
(c)

Figure 34. Scatterplots of the correlations between (a) fixation durations and numbers, (b) fixation durations and transition numbers, and (c) fixation numbers and transition numbers for experienced users on the location search task (app's text bar)

- Alert Settings

Figure 35 shows that there were very strong positive correlations among all correlation combinations of eye tracking variables on the alert settings task for first-time users.





(c)

Figure 35. Scatterplots of the correlations between (a) fixation durations and numbers, (b) fixation durations and transition numbers, and (c) fixation numbers and transition numbers for first-time users on the alert settings task

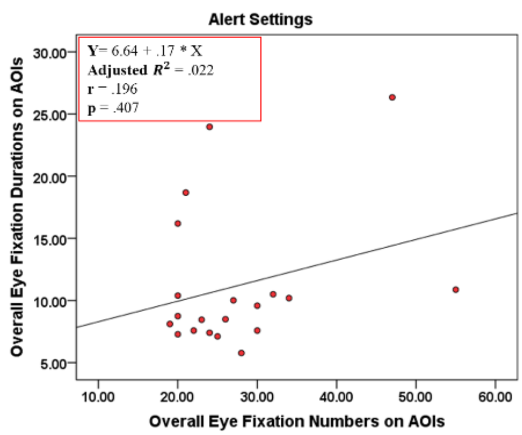


Figure 36. scatterplot of the correlation between fixation duration and fixation numbers for experienced users on the alert settings task

Figure 36 shows that there was a weak correlation between “fixation durations and numbers” on the alert settings task for experienced users. The correlations between transition numbers and both of fixation durations and numbers were not performed as all experienced users had the same number of transitions among AOIs on the alert settings task.

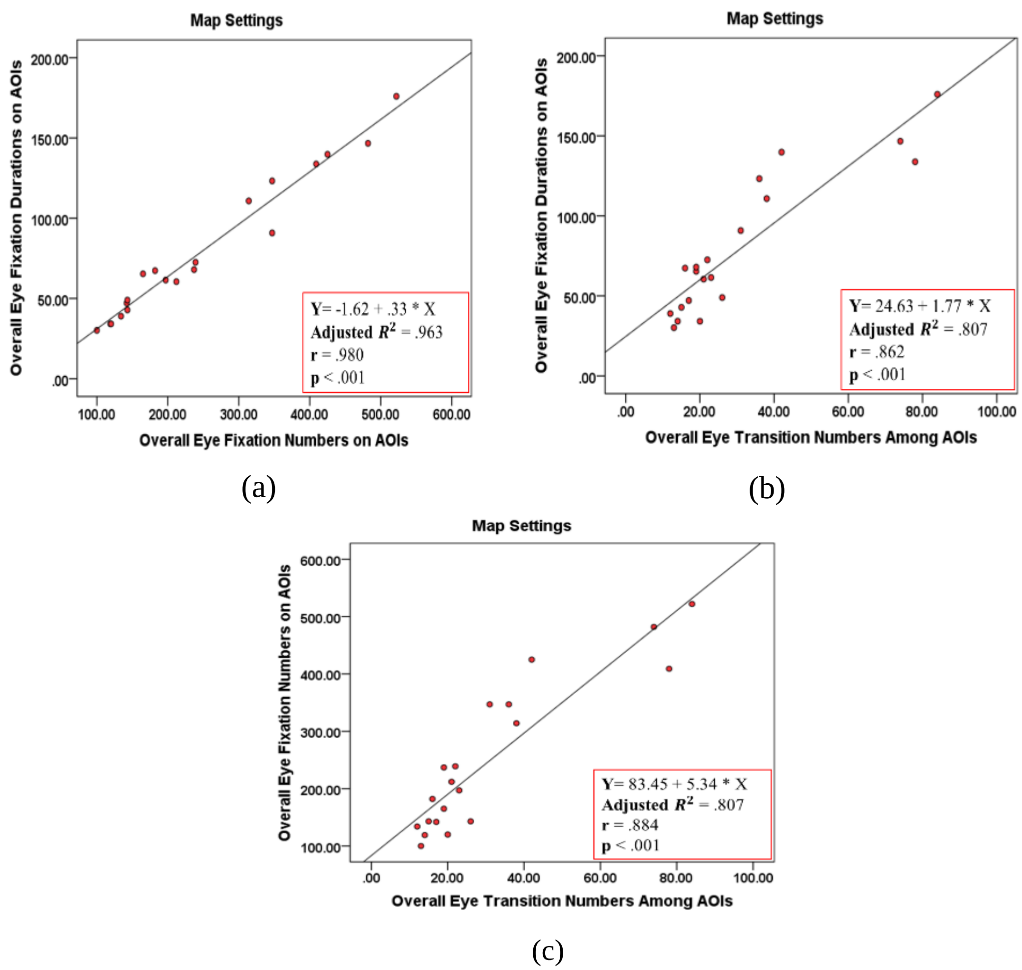


Figure 37. Scatterplots of the correlations between (a) fixation durations and numbers, (b) fixation durations and transition numbers, and (c) fixation numbers and transition numbers for first-time users on the map settings task

As shown in figure 37, there were very strong positive correlations among all combinations of eye tracking variables for first-time users on the map settings task.

Figure 38 shows that there was a weak correlation between "fixation durations and numbers" for experienced users on the map settings task. Similar to the alert settings task, the correlations between transition numbers and both of fixation durations and numbers were not performed as all experienced users had the same number of transitions among AOIs on the map settings task.

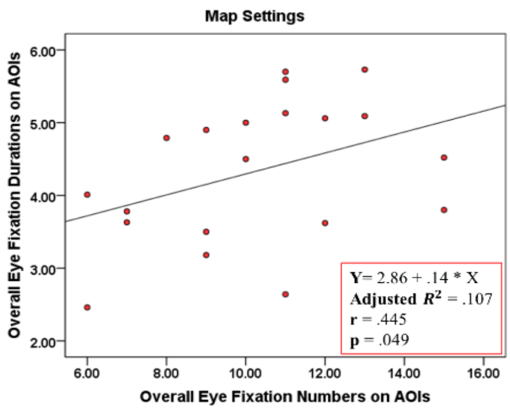


Figure 38. scatterplot of the correlation between fixation duration and fixation numbers for experienced users on the map settings task

Overall, the results for first-time users revealed that there were strong positive correlations between all pairs of eye tracking metrics for all tasks, indicating that increases in one metric were correlated with increases in the other metric. On the other hand, the results for experienced users showed that most of the data were weakly correlated and only three pairs of data were strongly correlated. However, the correlation coefficients for those pairs were barely above 0.5.

3.3.4 Discussion of Study II

The present study supports and confirms the findings from study I. In summary, we collected both first-time and experienced users' eye movement data while interacting with three main navigational tasks (location search, alert settings, and map settings), using one of the most widely used weather alert apps, the Weather Radio app. Through categorizing time-ordered and/or aggregated eye tracking data, we were able to better support the analysis and prediction of the users' cognitive processes. Overall, the study found substantially different interrogations of the displayed information between the first-time and experienced users for the alert settings and map settings tasks. The results imply that there were significantly different cognitive processes and scanning strategies between the two user groups on those two tasks (Bojko and Schumacher, 2008), as explained by their major difference of the overall eye fixation durations and numbers (see Figures 17 & 18). However, the two user groups appeared to have relatively similar cognitive processes on both approaches of the location search task. The analyses showed similar overall eye fixation durations and numbers for both groups on the two location search approaches, with significantly better performance on the app's text bar approach (see Figure 16). Several issues associated with usability principles, as well as correlation variables, are thoroughly discussed as follows.

The study found some affordance issues, which may have confused the first-time users. For example, the symbols, including the (i) symbol, located in the Bottom Menu "BM" AOI appeared to be counter-intuitive for the first-time users when performing the map settings task. The scanpath observations showed that most first-time users did not seem to expect the (i) symbol to be the function needed to reach the map settings menu; they seemed to randomly scan and tap the available symbols, trying to explore their functionality. The confusion about the functionality of those symbols is supported by the specific AOI analysis as substantially longer median eye fixations duration and larger median eye fixations number occurred on the AOI "BM" by the first-time users, as compared to the experienced users (see Figure 21). Another example of affordance issue was associated with the red pin symbol located in the Main Display "MD" AOI on the location search task. Controlling the pin on the map seemed problematic to most of the first-time users as they could not move the pin on their first try. Even though information regarding this matter was provided at the bottom of the AOI “MD” (Drag the pin to adjust your location), there was no indication that the pin needed to be tapped and held for over a second to be dragged. Although no significant difference between the two user groups was observed on the pin on map approach, the specific AOIs analysis showed that the first-time users had a larger number of eye fixations and longer duration of eye fixations on the AOI “MD” than those for the experienced users. These results suggest that interface symbols should be intuitively designed and/or associated with proper contextual information to ease users’ thought processes and clearly indicate their functionality (Gove, 2016; Norman, 2013).

The lack of visibility of some of the app’s functions could be one of the main factors for the first-time users’ deficient performance, especially on the map settings task. Specifically, the required symbol for directing users to the map settings menu when tapped, (i), was placed in an extended menu in the AOI “BM” that could not be seen unless a small map portion in the AOI “MD” was tapped (see Figure 7 (b)). The invisible menu seemed to significantly confuse the first-time users about the first step (finding the menu of the (i) symbol) needed for accomplishing the map settings task. This is supported by their complex and random eye transitions among AOIs (see Figure 30). The first-time users’ scanpath observations further supported this claim, as they revealed that most of the unnecessary, back and forth, and random AOI-transitions occurred at the beginning of this task, where they encountered the invisible menu. Even though smart phone app developers are restricted with the small screen size, it is critical for users' interactions with interfaces to prioritize usability by making menus and symbols clearly visible (Norman, 2013; Nielsen, 1994b).

The analysis also showed that the ambiguity of the expression "NWS" located in the AOI "MD" might be the reason for the first-time users' slow cognitive processing during the alert settings task. NWS stands for National Weather Service Organization and is the required option to direct users to the alert settings menu. The claim about slower cognitive processes is supported by the substantially longer eye fixation durations and larger eye fixation numbers on the AOI "MD" for first-time users, as compared to that for the experienced users on the alert settings task (see Figure 19). The eye transitions result further supports this claim, as most of the unnecessary AOI-transitions performed by the first-time users started from the AOI "MD" (see Figure 28 (a)). The figure also shows that the first-time users performed multiple back and forth transitions between the AOIs "BM" and "MD". This implies that once they visually scanned the AOI "MD" of the settings menu after tapping the gear symbol from the AOI "BM", they did not expect any of the listed options to be the required option for accessing the looked-for alerts; therefore, they shifted their eyes to other AOIs (mostly back to the AOI "BM") searching for and trying other possible options. Interface designers should abstain from using technical terms or jargon, such as the "NWS" in the settings menu, because they may affect users' understanding. Designers should rather use words and phrases that are easy to understand and familiar to the users (Nielsen, 1994b; Griffiths, 2015).

In terms of efficiency of use, the study showed that the pin on map approach of the location search task might not be an efficient option to search for and add a location. Specifically, the pin on map approach required several navigational steps to successfully accomplish the location search task (see Section 3.1). Each step was differently located on the smart phone display, which demanded multiple AOI-transition activities (see Figure 24) and a huge amount of overall eye fixation durations and eye fixation numbers (see Figure 16). In addition, due to the limited screen size of the smart phone, users were required to frequently zoom in and out and continuously scan the map to find the desired location. Searching the map seemed to impact both user groups, as most of their overall eye fixation durations and numbers occurred on the AOI "MD", which includes the map (see Figure 19). These results suggest the need for considering an efficient location search approach, such as the app's text bar approach tested in this study. We believe using such an approach will result in significant reduction in users' cognitive load (represented by their eye fixation durations and numbers), as was proved in this study.

The linear regression results showed that eye tracking data can be used to predict whether or not users are struggling when interacting with an interface. The results revealed detailed specific trends of how the two user groups' performances might be different. Specifically, the results (see Figures 31 to 38) illustrated that the correlations between the eye tracking metrics for first-time users' data appeared to be stronger than those for experienced users' data on all tasks with larger strength differences on the alert settings and map settings tasks. This implies that with the complexity of the app's interface, the first-time users appeared to uniformly focus on various parts of the display. In other words, the first-time users appeared to make numerous fixations on different AOIs with similar fixation durations and numbers of transitions among AOIs when searching for relative information. This explains the strong positive correlations among the first-time users' data. On the other hand, the weak positive correlations among the experienced users' data could be attributed to the experienced users' familiarity with the app prior to the experiment, as they seemed to pay more attention to the required elements for their tasks. The experienced users appeared to make a small number of fixations with different durations of fixations and numbers of transitions among AOIs, depending on the location and the relevance of information. In summary, the results indicate that if users find difficulty in processing and extracting the needed information from a particular expected location, they tend to overall search and scan the display with poor efficiency.

3.4 Chapter Summary

This chapter has presented two studies with the aim of identifying usability issues that result from end-users' interaction with smart phone weather apps. In addition to identifying the issues, we aimed to figure out the reasons for committing those issues and their impact on users. Moreover, we proposed alternative approaches for some of the tasks to get an idea if they would enhance users' experience, compared to the existing features. We employed two evaluation methods: traditional method (completion rate, task time, and user satisfaction) and eye tracking method (eye fixation duration, eye fixation number, and scanpath patterns). Our main focus was on first-time users, as usability issues are commonly found from first-time users' interaction with interfaces (Gerardo, 2017). However, we included experienced users to serve as a reference measure and explain how user-friendly the interface is for all users, regardless of their experience levels.

In Study I, we tested four commonly used features on weather apps: location search, alert settings, map settings, and alert messages. Even though all users successfully completed the given tasks, the results showed that first-time users had substantially longer completion time than experienced users on both the alert settings and map settings tasks. For the location search task, both user groups had similar completion time on both location search approaches. However, both groups completed this task using the app's text bar approach with a significantly shorter completion time than that with the pin on map approach. For the survey comparisons of weather alert messages task, both groups rated the given survey items similarly on both message examples. However, they were highly satisfied with the proposed messages and were extremely disappointed with the existing NWS messages.

Several usability issues were identified, based on both the objective and subjective results, such as lack of visibility, lack of affordance, inefficiency, and poor use of language. For example, the map settings task included invisible menus that required previous operational steps. Specifically, to access the hidden map settings menu, a user needs to tap a tiny map part at the bottom of the screen. This visibility issue was coupled with another issue, which was that after finding the map settings menu, there were no labels or intuitive icons to indicate the map settings' functionality.

Study II supported the findings from Study I and informed users' thinking processes and objectively provided more details about the discovered issues. Overall, significantly larger eye fixation numbers, longer eye fixation durations, and complex scanpath patterns were observed for first-time users, compared to experienced users. The results indicate that first-time users struggled finding the needed information for their tasks as well as processing that information once they found them. As an example about the added value of eye tracking metrics, the first-time users' poor performance on the alert settings task, was mainly due to the terminology used in the AOI "MD". Significantly larger eye fixation number and longer eye fixation duration were observed on that AOI for first-time users, compared to those on the same AOI for experienced users. In addition, it was found that several unnecessary eye transitions started from the AOI "MD". These results indicate that once users sufficiently scanned the AOI "MD" and did not expect any option to be the needed to access alert settings, they scanned other AOIs searching for possible needed information.

Chapter 4. User-Centered Design Assessment of End-User Needs in Smart Phone Weather Applications (Phase 1)

Author's Note: The content in Section 4.2 will be submitted soon (Status: ready for submission) as a conference paper to the 2019 Human Factors and Ergonomic Society (HFES) Conference. The author of this dissertation wrote the paper in collaboration with his advisor, Dr. Ziho Kang, and his previous lab mate, Dr. Elizabeth Argyle.

4.1 Introduction

This chapter presents the first phase of the user-centered design process: investigating user characteristics, issues, goals and particularly their needs in interface designs, as shown in Figure 2. As illustrated in Chapter 1 and 2, it is critical to design interfaces from the user perspective. In the first part of this chapter, we perform focus group sessions with regular smart phone users of different weather apps to learn about their experiences with existing apps and how their experiences can be enhanced in future designs. Specifically, with the knowledge acquired from the findings of the studies in Chapter 3, we present critical questions to the participants and encourage them to share and discuss their opinions and suggestions in a group setting. The focus group findings are believed to greatly help developers and designers know about their intended users' exact requirements so that they consider them in upcoming designs.

In the second part of this chapter, we present sets of essential usability guidelines and heuristic specifications that are commonly used for smart phone app designs as well as widely used in the literature. In addition, as discussed in Chapter 2 that older users have special limitations and characteristics, we also show several heuristics that have been proven in the literature to enhance older users' experience and confidence with smart phone apps.

Overlooking user needs and different characteristics in the design stage may prevent intended users from effectively, efficiently, and/or comfortably using an interface system and achieving its main goals. Systems with poor usability, especially those involving time-critical data, such as weather data, can be associated with several system errors, and/or slow response time, which can hinder performing life-saving actions.

4.2 Users' Perceptions of Smart Phone Weather Applications' Usability: A Descriptive Qualitative Assessment

Several UCD methods are used and reported in the literature, such as usability testing (Feng et al., 2017; Argyle et al., 2017), questionnaires (Bias et al., 2012), and focus groups (Schnall et al., 2016; Argyle et al., 2015). Specifically, focus groups are among the most common UCD methods and can provide important qualitative insights into system designs. Nielsen (1997) defines focus groups as a semi-informal procedure involving a group of participants (6-9 people in each focus group) to discuss a certain topic in a structured manner. Focus groups are commonly used in the human factors field, as they are capable of providing detailed qualitative information about a fundamental theme. They are typically guided by an experienced moderator who leads and directs the discussion based on the objective of the study. Moderators must have sufficient experience with the subject of interest as well as with guiding the discussion, in an unbiased manner, to attain the intended results (Caplan, 1990). Focus groups help designers elicit actual users' feelings towards, issues with, and requirements for systems at both early and late interface design stages.

4.2.1 Method

Participants

Fifteen participants (9 male, 6 female) were recruited to participate in the focus groups. Two focus groups were conducted, and each lasted for about 90 minutes; the first focus group included seven participants and the second one had eight participants. The number of participants in each session considered in this study was based on Nielsen's recommendation (Nielsen, 1997). The mean age of the participants was 32 years (*SD* = 6.4). The participants were recruited in three different ways: (1) a mass email sent to University of Oklahoma students, (2) flyers taped to the doors of multiple public buildings, and (3) personal communication with friends and colleagues. All participants were users of popular smart phone weather apps that include typical features (e.g. location search, weather forecasts, radar/map, and alert notifications/messages); none of the participants were experts in meteorology and/or used advanced or technical weather apps. In addition, at the time of the experiment, the participants actively used eleven different weather apps, run on various operating systems, with more than two years of usage.

Focus Group Design

The focus groups were guided by a skilled moderator, who had large experience with the subject matter and with discussion leadership. The moderator asked the main questions followed by probing questions to completely understand participants' opinions, suggestions, and concerns. In addition to the moderator, an assistant to the moderator took part in leading the discussion. The assistant to the moderator took notes of all participants' feedback and gave a summary of the discussion at the end of each session.

The focus groups addressed a variety of topics related to the usability of smart phone weather apps and their interface designs. The questions meant to elicit participants' feelings towards the general usability of existing weather apps, the issues with specific weather app features, as well as their views on future interface designs of great usability. Specifically, the questions were structured as follows:

- General information (4 questions)
    - Tendency for downloading particular weather apps
    - Priority of features on weather apps during both time-critical and non-time critical weather conditions
    - Positive & negative usability experiences with weather apps
    - Comments on the discussed matters
- Specific information (5 questions)
    - Use of exact and familiar location feature
    - Control of alert settings
    - Use of descriptive information
    - Presentation of pushed alert messages
    - Use of menu icons & labels

A full list of the questions given to the participants can be found in Appendix C. The questions were intended to produce open-ended answers and the participants were encouraged to make follow-up comments/responses to enhance the discussion and outcomes.

Procedure & Data Collection

Prior to conducting the official focus group sessions, a pre-test mock-up focus group session was conducted to examine the validity of the questions and become aware of any issues to consider during the official sessions.

The official focus groups were conducted at the University of Oklahoma, Norman campus, in a controlled environment. The participants signed consent forms and filled out a demographic survey upon their arrival. Following that, they were introduced to the objective of the study and encouraged to actively participate in the discussions. Then, the discussion began. The focus group questions were projected on a large whiteboard for participants to reference during the discussion. The focus groups were video recorded to capture all participants' answers and reactions.

Data Analysis

The authors transcribed the recordings verbatim and then analyzed them using the thematic analysis approach (Boyatzis, 1998; Braun & Clarke, 2006). Thematic analysis is a qualitative data analysis method that examines the dataset to identify themes (Braun & Clarke, 2006). For the analysis, the whole dataset was coded based on the objective of the study and then the codes were categorized into themes and sub themes. The codes are words or short phrases that summarize and/or paraphrase ideas and feelings stated in the data (Elizabeth, 2015). The themes and sub themes are created based on those codes that are further defined by Saldana (2015) as a short word or phrase that symbolically assigns a summative, salient, essence- capturing, and/or evocative attribute for a portion of language-based or visual data."

Thematic analysis consists of four iterative steps: (1) preliminary coding, (2) code modification, (3) creating themes and sub-themes out of codes, and (4) structural validation (Braun et al., 2014). This study majorly follows the procedural framework introduced by Braun and Clarke (2006), which contains a step-by-step guide to performing a thematic analysis. The steps are summarized in Table 6. In thematic analysis, we deeply investigate and critically think and interpret participants' responses in order to produce reliable conclusions. In particular, we summarize themes and examine the relationship among them using our interpretation in a way that complies to our main goal; this process is called semantic approach (Braun & Clarke, 2006).

Table 6. Steps for performing a thematic analysis, summarized from Braun and Clarke (2006)

Phase	Activities
(1) Self-Familiarization	Transcription and investigation of data (transcripts, media, etc.)
(2) Initial Coding	Generation of codes and patterns in the data
(3) Searching for Themes	Grouping and categorizing codes under themes
(4) Reviewing Themes	Refinement of themes
(5) Defining Themes	Explanation of themes with respect to the overall research goal
(6) Reporting Output	Description of themes with representative examples from the collected data

4.2.2 Results

In the current study, we analyzed the data in terms of the generated themes as well as the frequency of occurrence. The frequency of occurrence is used here to indicate the significance of the themes and their impact on the participants.

Four major themes related to usability principles emerged from the thematic analysis and further captured a total of 10 sub themes, shown in Table 7. The four main themes were user cognitive load, effectiveness, efficiency of use, and user perceptions.

Table 7. Focus group findings (themes, sub-themes, and frequency of occurrence)

Themes	Sub-themes	Frequency of Occurrence
User Cognitive Load	Affordance	45
	Amount of information	39
	Use of language	19
	Information visualization	16
Effectiveness	Priority of information	44
	Ease of use	33
	Flexibility in use	9
Efficiency of Use	Task time	35
	Number of operational steps	24
User Perceptions	User perceptions	13

- **User cognitive load**

The participants described a need for weather apps that require minimal cognitive load. In the context of user experience with system interfaces, cognitive load refers to the total mental effort a user expends in searching for and processing information on a display to execute a desired action (Whitenton, 2013).

The participants seemed to be mostly concerned with the intuitiveness of smart phone weather interfaces. They expressed their desire to interact with weather apps that work the way they expect. They also shared examples of confusion when using features built into their current weather apps. One participant stated:

> "Even though my weather app is very accurate, using some features like radar is annoying, especially when I leave the app for some time. I almost always forget how to access the radar options as they are placed in [an] unexpected location. Also, their icons do not indicate what they are really for."

In order to use weather apps that comply with users' mental models, the participants agreed that developers should place functions in anticipated locations and use representative icons that are understood by users of different age groups and experience levels. In particular, to enhance the intuitiveness of the function icons, one participant suggested and agreed by others:

> "Weather apps should use more standardized icons that are used across every app, or at least label them to conveniently [intuitively] indicate their functionality."

Many participants expressed their dissatisfaction with the large amount of information shown through their weather apps. The participants thought that with the time-critical situation associated with some weather conditions (e.g. tornados and floods), users would struggle finding and processing the needed information if the interface was full of information. In reaction to a sample of flood warning alert message (pushed to weather app users in Cambridge, MA on Apr 1st, 2017) shown to the participants (see Appendix C), one participant stated:

> “What do I look for? There is too much going on here! How come all that information displayed on the small smart phone screen?”

Some of the participants noted that most of the presented information was unnecessary and/or irrelevant to the users’ saved location. For example, they thought they did not need to receive a list of all areas under the issued alert; they believed including information about only the user saved location would be needed.

Another factor discussed by the participants was the use of language throughout weather apps. The participants named and criticized several weather apps that use weather-expert-level terminology and/or jargon that is not understood by everyday users. For instance, the participants commented that they had experienced receiving alerts that included numbers and codes that they could not understand. Those codes and numbers represent the areas under the issued alerts, which are included in the raw alerts weather app operators receive from local agencies, such as National Weather Service (NWS) centers. The messages are not interpreted by the weather app operators; they are pushed directly to end-users.

Several users mentioned that the way weather information is visualized would influence their decision to download and use smart phone weather apps. In the words of one participant:

> “I care a lot about visualization. I prefer apps that use indicative images, so I can feel what is really going on outside, like rainy image when there is rain, not only text.”

In addition, some participants stated that organizing the weather information in a hierarchical manner (e.g. first showing current weather forecasts, then hourly forecasts, and finally daily forecasts within the same screen) would greatly ease their mental process in locating the required information.

- **Effectiveness**

The participants discussed the importance of interacting with effective weather apps. Effectiveness refers to the extent to which users are able to achieve a desired goal with ease, flexibility, and accuracy (Shackel, 1991).

The participants repeatedly expressed that optimizing weather app interfaces by focusing on the most important feature would greatly enhance their experience. However, they did not agree on a single feature to be the most important. Most participants stated that the daily and hourly weather forecasts are extremely important because they are frequently viewed. One participant mentioned:

> "I believe weather forecasts are the most important information to look for… because most of us [users] regularly check this information and that is sort of the main reason why we would go to a weather app."

Some participants thought the alert messages were of greater importance because of the life-threatening risk associated with the embedding weather conditions. Other participants believed that a location search feature should be prioritized due to its need during travel.

The ease of using weather apps was one of the participants' main focuses during the discussions. The participants agreed on the importance of having a user-friendly interface which enables a user to easily navigate the app and always performs tasks in a straightforward manner. As a violation to the ease of use principle, one participant shared an example regarding a weather app that restricted users search for locations by only navigating an embedded map to find and select the desired location:

> "It is very difficult to find a location on the map considering the limited screen size of the phone and maybe the lack of familiarity with the geographical area… not all users can do that."

Some of the participants pointed out the need for performing tasks with some sort of flexibility. Several ideas were provided. One of the ideas was to include shortcut alternatives in addition to the traditional ways of performing tasks. The participants attributed the need for shortcuts to instances of time-critical weather conditions that require fast access to relevant information. Another idea was to enable users to search for locations in multiple ways including location name, address, zip codes, or city and state. The participants believed that having this flexibility during searches would accommodate a wide variety of users and improve the overall experience.

- **Efficiency of use**

Highly efficient systems are the ones that require the minimum inputs to attain the maximum outputs (Shackel, 1991). Task completion time and number of steps to complete a task on a system are among the most popular indicators of systems' efficiency (Albert & Tullis, 2013).

The participants emphasized that the time to complete any task on any interface should be minimal and that weather interfaces must be among the most efficient interfaces. When the participants were asked about their preferences of using exact location by typing the location name vs. searching for the location on the map, one participant stated his opinion and agreed by the others:

> "Typing the exact location in the search bar with auto suggestions will save me so much time. The developers should think from the user's side."

The participants also agreed that a limited number of operational steps to complete any task would highly enhance user experience. One participant commented:

> "I don't see any point of having to go through multiple navigational steps to access critical information such as weather information, especially on [a] smart phone platform. Once developers consider what the user really needs, they can have everything within one or two screens."

- **User Perceptions**

User perceptions refer to how a user recognizes, understands, and interprets a certain element (Encarnaçao et al., 1994). Some participants responded to questions based on their perceptions of what was asked. For example, in a response to the participants' preference of the use of the location search feature, one participant stated:

> "From my experience, especially here in [my hometown], it's better to add your house address. This is because apps show you that it's raining in school and you go home and it's not that much rain. So, I prefer to use my exact location because I will feel more confident about the accuracy of weather information."

4.2.3 Discussion

Major Findings

This study investigated users' opinions towards the usability of existing smart phone weather apps and identified their requirements for future usable apps. The results indicate that the participants were mostly concerned about the risk associated with imminent weather conditions.

Specifically, the focus groups produced several usability issues in current weather apps, as well as requirements for future designs. The usability issues and requirements are characterized in the form of themes and sub-themes. Most of the emergent themes and sub-themes align with the usability principles for interface design introduced by Nielsen (1995b) and discussed by Rogers et al. (2011), suggesting the need to consider such principles in the design of all system interfaces. The usability principles are explained in detail in the next section of this chapter. The findings further suggest that the current smart phone weather app market has room to develop and meet the dynamic needs of its population of users.

The participants' frequent call for intuitive design, less information, an appropriate use of language, and nice visualizations, implies that it is critical for weather app interfaces to lessen users' cognitive load: to ease finding and processing the presented information. This is supported by the participants who voiced concerns about the hazard levels associated with some extreme weather conditions that require prompt and appropriate reaction. For example, the participants stated that the weather alert messages pushed to their devices were not user-friendly due to the inclusion of technical terms, the large amount of information in messages, and the irrelevant information to their exact affected location. Even though these factors are expected to substantially help users find and comprehend the relevant content of the alert messages, limiting information to the user's exact location may make some users miss out on critical information in some specific contexts of use (e.g. when traveling). The effectiveness of the alert messages and their perceptions by users are currently under profound investigation by NWS experts (Jacks et al., 2018).

Despite the disagreement among the participants at the beginning of the discussion on the most important feature in weather apps, most of them later agreed that both the alert message and weather forecast features were of the same importance. One of the participants mentioned:

> "If an app prioritizes these two features [pushed alert messages and weather forecasts] and makes them easily accessed and understood, I would definitely use it."

The participants also emphasized that weather apps should be optimized in a way that enable users to complete any task with the least amount of time and number of operational steps. This finding is in line with one of the findings in Kaufman's (2016) study that most smart phone app users seek to spend as little time on a task as possible, so they can quickly resume their daily activities.

4.3 Usability Heuristics for Smart Phone Interface Designs

After identifying users' specific concerns with current smart phone weather apps and needs for future usable ones, it is very crucial for designers and developers in the

UCD process to synchronize those specific needs with essential usability heuristic and smart phone design guidelines. This synchronization is believed to produce highly usable interfaces (Manzari and Trinidad-Christensen, 2006). It is worth reminding that the focus group findings were largely in line with several widely accepted and commonly used usability heuristics. Usability heuristics are a set of guidelines to be followed for usable interfaces (Nielsen, 1995b).

4.3.1 General Essential Usability Heuristics

There are thousands of usability heuristics available in the literature. However, that large number of heuristics is usually perceived as intimidating by usability practitioners and interface developers (Nielsen & Molich, 1990). In 1989, Nielsen and Molich created a list of nine usability heuristics. Those heuristics were developed based on usability problems found in basic telephone and computer devices. Even though the heuristics were created based on problems found in only one basic interface, they are still applicable to be applied for many other advanced interfaces. However, with the revolution of technology and proliferation of smart phone devices and touch screen interfaces, a call for updated usability heuristics was raised by several authors.

In 1994a, Nielsen revised the previously published usability heuristics. His revision included benchmarking 249 usability problems found in 11 popular products of multiple characteristics to 7 previously published sets of general usability heuristics. He categorized each of his revised sets under representative names that clearly described the underlying usability characteristics. The 7 usability heuristics by Nielsen (1994a), along with their definitions are as follows:

1. Visibility of system status

A user should always and immediately be aware of the system status, meaning that any action is performed on a system by a user, should be followed by a proper feedback in a short time.

2. Match between system and the real world

It is very important for any system to be a user-oriented rather than a system-oriented, meaning that the language used in the system (e.g. words, expressions, and notations) should be familiar to intended users, used in their everyday lives, and follow a natural and logical order.

3. User control and freedom

As any system interface is developed for users, it is necessary to provide them with high levels of control and freedom when using system's features. In addition, a system should make it easy for users to undo and redo their actions, as users sometimes do mistakes or perform undesired actions. Users must not pay the price (by having to go through several steps) to recover from mistakes.

4. Consistency and standards

This heuristic is considered one of the most important usability heuristics in interface designs. A system functions, menus, dialogue boxes, layout, and icons of the same characteristics should be consistent throughout the entire interface so that users can intuitively understand the presented information and react accordingly.

5. Error prevention

A system developers should first carefully examine their interface design, perhaps through performing pilot studies with target users, and anticipate the errors that might results from users' interaction with their systems and try to prevent them. However, errors are inevitable and highly expected. Hence, a system should provide users with clear and concise error messages.

6. Recognition rather than recall

It is very important to make all system's elements visible for users so that users can easily distinguish the needed information and intuitively execute actions. Users should not extensively use their memory to retain information that leads to another information within the system.

7. Flexibility and efficiency of use

A system should be flexible in a way that meets the needs of different users' experience levels with the system. For example, instead of typing a location in the search bar every time a user uses the system, returning users should get an advantage over first-time users by having an option to access the previously searched-for locations to minimize time and effort.

Nielsen (1994a) also added two more usability heuristics as follows:

8. Aesthetic and minimalist design

A system developer should avoid any unnecessary or infrequently needed information and rather prioritize the information that their end users always need to perform any task. This is emphasized due to the fact that any additional irrelevant information increases users' cognitive load and therefore hinder their performance and affect their satisfaction.

9. Help users recognize, diagnose and recover from errors

Errors should be easily communicated to users using easy language and including actionable recommendations for recovery in a very precise manner.

Nielsen and Mack (1995b) later added one last heuristic as follows:

10. Help and documentation

If a user struggles on an interface and requires help, then there is something wrong with the interface. A usable interface lets the user interact with its features naturally and without any training or help. Although it's recommended not to include any documentation, it may be needed in some interfaces of complex features in nature.

The aforementioned ten usability heuristics are the most widely known and commonly used heuristics that are applicable to any interface design (Douglas, 2017). Even though Nielsen's ten usability heuristics are crucial for interfaces on different platforms, including the smart phone platform, they are either too broad or too general and do not sufficiently suit users of special characteristics, such as older users (Silva and Holden, 2014). In addition, this general list of heuristics might not provide a comprehensive and specific guide for designing usable smart phone apps. Therefore, in the next two sections, we discuss more usability heuristics that are applicable to the goals of this dissertation. Specifically, in section 4.3.2, we show a comprehensive list of heuristics used in developing smart phone interfaces based on the needs of older users. In section 4.3.3, we present a list of specific smart phone app design guidelines that are published by reliable and popular sources, such as Google and Apple corporations.

4.3.2 Specific Usability Heuristics for Smart Phone Older Users

As indicated in Chapter 2, older users face several limitations that arise naturally with aging. Among the most common age-related limitation are declines in vision, hearing, motor skills, and cognitive abilities (Ilyas, 2010). Vision issues include eyesight deterioration, difficulty in perceiving and distinguishing certain colors, and reduction in pupil size (Silva and Holden, 2014). Hearing problems can be either gradual hearing loss or sudden complete hearing loss (Silva and Holden, 2014). Motor skills decline may lead to slower reaction time and reduction in flexible movement (Fisk et al., 2009). People's cognitive abilities are also largely diminished with age. The most affected cognitive ability among older users of smart phones is working memory (Fisk et al., 2009). Examples of weakened working memory in smart phone use include failure to navigate complex menus, cognitively process displayed information, and recall operational steps to complete a desired task (Fisk et al., 2009). Loss of attention, language comprehension deterioration, and memory loss are also signs of deterioration in cognitive abilities. As our present work does not include audio functionality, we are neglecting hearing problems and related recommendations and heuristics.

Numerous usability heuristics exist in the literature to accommodate the needs of older users in systems' interfaces, including those of smart phones. The heuristics were derived from both qualitative and quantitative studies with older people. It is worth stating that these heuristics are not only applicable to older users, but also to younger users; they are associated with older users because they were found to greatly boost older users' confidence, satisfaction, and performance. For instance, Hawthorn (2000) conducted a literature review of 100+ studies on technology and age-related factors and ended up providing several heuristics and guidelines for usable smart phone interface designs.

Visual design heuristics

The visual aspects of interface designs, especially smart phone interface designs, are among the most important design aspects, as the information are mostly communicated visually (Pak & MacLauglin, 2010). The visual heuristics include simple and clear layout that help users to easily finding and processing the looked-for elements (Khawaji, 2017). To achieve that, it is recommended to use bright screen, relatively large icons and font size (e.g. 12 to 14 point font), limited number of colors, group related elements, large spaces between items, and appropriate color contrast (e.g. black text on white background) to enhance readability (Silva and Holden, 2014; Kurniawan and Zaphiris, 2005). In addition, it is highly recommended to avoid using any moving objects, flashing elements, and information on peripheral screen areas (Khawaji, 2017).

Heuristics for cognitive and Motor aspects

To account for older adults' cognitive and motor issues, it is recommended to minimally delay processes, use simple and intuitive designs that match with users' mental models, and recognize objects and steps and not to retain them in memory (Hawthorn, 2000). In other words, it is extremely important to design simple menus with limited necessary options, label icons with appropriate text (Barros et al., 2014; Sharma et al., 2012), and avoid clutter of text and graphics (Silva and Holden, 2014; Kurniawan and Zaphiris, 2005). In addition, Chisnell and Reddish (2006) found that avoiding jargon and technical terminologies would highly enhance older users' experiences.

4.3.3 Smart Phone Application Design Heuristics

The smartphone app market is getting bigger so fast. Around two billion smart phone apps were downloaded in 2017 alone (Statista, 2018c). With the great competition among app developers and the availability of numerous apps of the same services, users tend to target those that are user-friendly. Statistics showed that nearly 25% of app users install and open an app once and never use it again (eMarketer, 2015). The vast majority of users aim to finish an app's tasks as efficient as possible so that they return to their normal day activities (Kaufman, 2016). This reminds us of the famous saying by Steve Jobs: "It's not just what it looks and feels like. Design is how it works." Nielsen (2009) also believed that any app interface should enable users to complete their tasks quickly.

In this section, we highlight some of the most important smart phone app heuristics. The following heuristics are iOS-based heuristics and are summarized from Kaufman, (2016), Gove, (2016), and Griffiths, (2015).

1. Display the app's value clearly and first

It is important to enable users to comprehend the main goal of the designed app as soon as they install and access the app. One way is to provide a very clear "call to action" function, centrally positioned, so that users can access the function right away.

2. Focus on the most important feature

App developers must know the single-most important and/or frequently used feature, so they can give it more priority in the design. This is expected to facilitate users' interaction and satisfy them.

3. Swiping screens

Scrolling or swiping screens is highly recommended for apps that include structured and focused content.

4. Organize and label menu categories to be user-friendly

Users may have difficulty in understanding and accessing app menu options that do not comply with their mental models. Hence, menu options should not be overlapped and should use easy and different words/phrases to be easily distinguished.

5. Provide helpful descriptive information

App developers should know that users have different experience levels and characteristics. If possible, provide additional simple descriptive information to clearly indicate the functionality of the app's features and menu options.

6. Enable users to "go back" easily

Users sometimes need to access the immediate previous screen when using the app. It is very important to enhance the flexibility of this functionality by enabling users to go one step back throughout the entire app. Forcing users to start over may lead to losing unsaved information.

7. Position elements appropriately

Each app element has an ideal location and size, depending on its importance level. It is crucial that app developers employ the mapping principle in this matter. For example, the back arrow should be placed on the left side to be mapped well with its action.

8. Make it easy to manually change location

For the apps that include location search feature, it very important to meet the users' different needs. The app should use the auto-detection functionality for efficiency purposes. In addition, users should be able to manually change locations that suit their needs. The manual location feature should be straightforward.

9. Use effective search indexing

This heuristic ties to the previous one. Apps with location search feature should use useful search index that is updated based on users' inputs. This is mentioned due to the fact that users will expect any location search index to work exactly as the popular Goggle map app. Example of effective search indexes include auto-suggestions, auto-corrections, and recently searched-for locations.

10. Provide text labels and visual keys to clarify visual information

As explained earlier, providing representative icons for the app's features as well as labeling them with appropriate text greatly help users in understanding the underlying functionality.

11. Use constant navigation menus

To speed features' accessibility and quickly compensate for mistakenly accessed features, it is recommended to keep the navigation menu fixed throughout the entire app.

4.4 Chapter Summary

This chapter has described the findings related to the first phase in the UCD process: user needs and limitations among smart phone weather apps. In addition, the most widely accepted general usability heuristics, specific heuristics for older users, and smart phone app design heuristics are described in this chapter.

Overall, the focus group findings showed that the participants were mostly concerned about the cognitive aspects of weather app designs, especially during time-critical weather conditions. Specifically, they called for intuitive designs, minimum necessary information, easy to understand content, and nice visualization. They attributed that to instances of extreme weather conditions that require quick information accessibility and reaction. They also discussed needs related to the effectiveness of weather apps, such as priority of features and displayed information, simplicity, and flexibility of use. Efficiency of use and perception of presented information were also discussed.

Nielsen's (1995a) usability heuristics as well as the described necessary guidelines for older users and the smart phone app design heuristics were largely in line with user needs voiced in the user group sessions. This indicates that carefully addressing those needs following the discussed heuristics in future designs will likely result in substantially enhanced user experience. In the next chapter, we present the second phase in the UCD process: developing a prototype based on user needs, where we show the whole structure and content of the developed weather app prototype.

Chapter 5. Development of A Smart Phone Prototype Weather Application Based on the User-Centered Design Approach (Phase 2)

5.1 Introduction

This chapter presents the second phase (phase II) in the UCD process, which is the development of a smart phone prototype app from the end-user perspective. In this prototype app, we aimed to address all users' issues with currently running weather apps and needs for future ones, which were discussed in the focus group sessions (see Chapter 4). All weather features (location search, weather forecasts, alert messages, map settings, and alert settings) embedded in any downloadable smart phone weather app were enhanced based on end-users' feedback. In addition, we carefully considered the age-related limitations and heuristic guidelines (illustrated in Chapter 4) to enhance older users' interaction with time-critical weather apps. Moreover, we employed the most popular and commonly used usability heuristics in the interface of this prototype app. This prototype app is intended to be used by both younger and older users.

5.2 Structure and Content of the Prototype Application

This smart phone prototype weather app was created by the author of this dissertation with guidance from his advisor and was called "EZ Weather". The development process included: 1) creating the interface screens using the Photoshop (https://www.photoshop.com/) software and then (2) developing the app using the InVision app software (https://www.invisionapp.com/company). InVision app is an online interactive software that is intended to create high-fidelity web and smart phone app prototypes; InVision is capable of developing smart phone prototype apps of different operating systems (e.g. iOS and Android); in this dissertation, we developed an iOS prototype app that ran on the iPhone 6. InVision software has very advanced and powerful algorithms, which help developers build prototypes that suit their needs in a very flexible manner. Using the software, developers can easily and flexibly add interaction and transition features that map well with the intended user actions.

The specific design of the weather prototype app (EZ Weather) took into consideration the user needs and multiple heuristic guidelines described in Chapter 4. The heuristic guidelines with visual examples from the developed prototype app are explained as follows:

1- Clear call to action

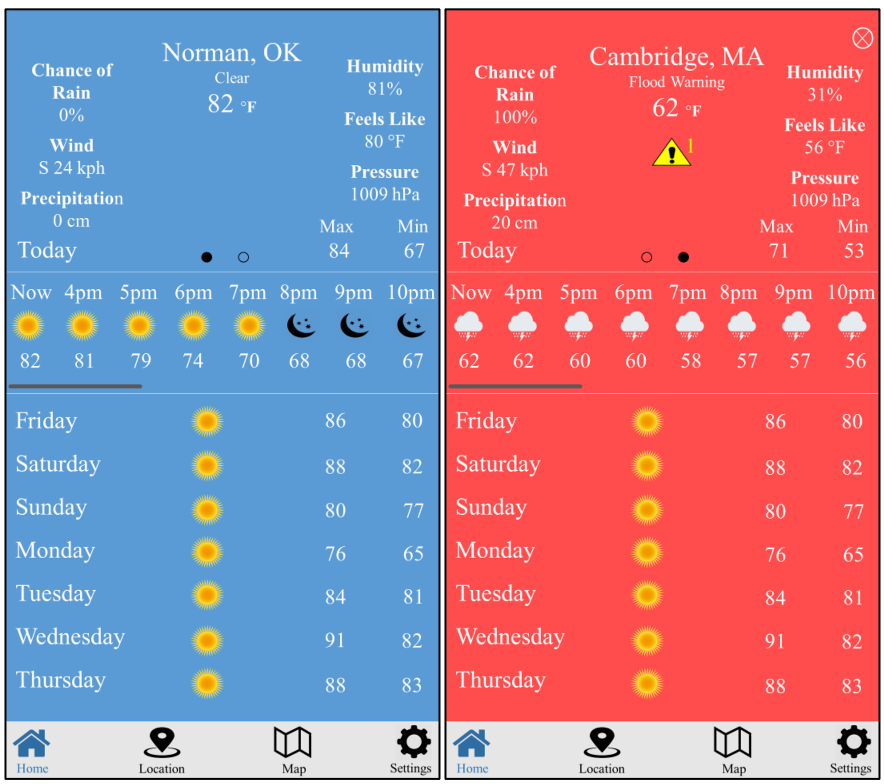
Figure 39. Example of clear call to action on installation screen

As soon as users install the app, they are interacted with a screen that provides a very clear and intuitive call to action. In the case of weather alert apps, the first task a user needs to perform is adding a location for which they can access weather forecasts. EZ Weather shows on the installation screen both auto-location detection or manual location entry fields for users' choice, depending on their needs (see Figure 39).

2- Optimization of the most important feature

3- Use of appropriate visualization with indicative colors

4- Swiping screens with visible and intuitive indication



(a) (b)

Figure 40. Example of optimized features with the use of appropriate visualization and swiping functionality

As voiced by the participants in the focus groups that both weather forecasts and alert messages features are the most important features and share the same importance; the choice of weather forecasts feature because of its frequent accessibility throughout the day and the alert messages feature because of the time-critical situation associated with extreme weather conditions. Hence, EZ Weather provides an all-inclusive weather forecasts feature, where all forecast information (daily, hourly, and extended current forecasts) are associated with their saved location within a single screen (see Figure 40 (a & b)). This single screen is used as a home screen; it is directly accessed as soon as a user opens the app. In addition, when a weather alert is issued and pushed to users, EZ Weather displays a representative alert icon (see yellow alert icon in Figure 40 (b)) on the affected location so that users can tap and access the related messages. Moreover, as shown in the figure, all information on EZ Weather are grouped, in terms of relevance, and nicely visualized. In case of a life-threatening weather warning alert at a saved location, not only information is presented but also an indicative red background color is used to intuitively and easily realize the matter (see Figure 40 (b)). A blue background color is also used to indicate normal weather condition (clear sky) (see Figure 40 (a)). The swiping functionality is also used in EZ Weather, where users swipe right or left in order to access different saved locations; carousels (little circles) are used to indicate the swiping feature, where the filled circle means current screen (see Figure 40 (a & b)).

5- Use of representative icons and appropriate text labels

6- Use of helpful descriptive information

7- Use of simple interface with limited necessary options

8- Generous spacing between items

Figure 41. Example of appropriate use of labeled icons and descriptive information

EZ Weather associated all menu options and bottom navigation bars with representative icons (e.g. "i" icon to show information about the app) and intuitive text labels (see Figure 41). In addition, the minimum necessary information and menu options were used throughout EZ Weather (e.g. 4 menu options in the general settings menu) to support the simplicity of the interface as shown in Figure 41. Furthermore, descriptive information was added to further help users fully comprehend the functionality of the underlying feature prior to tapping (see text under each menu option in Figure 41). EZ Weather also considered thumb ergonomics (Anthony, 2012) to account for the size differences in users thumbs and to avoid any slip errors (see spaces between icons at the bottom navigation bar in Figure 41).

9- Use of consistent design

10- Use of minimalist design

11- Use of effective "back" functionality

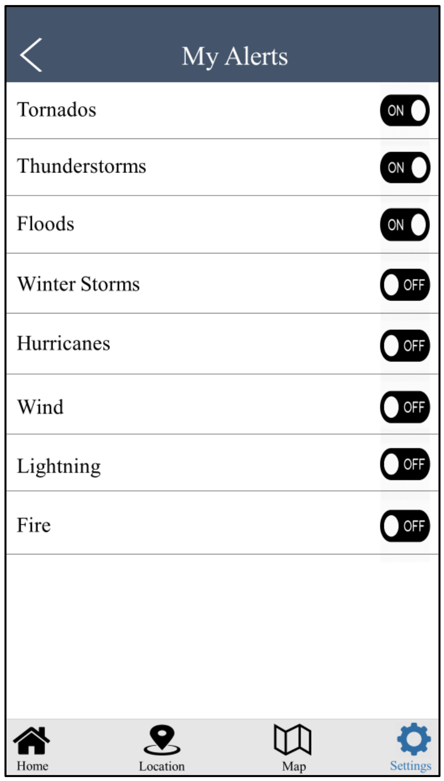


Figure 42. Example of consistent and minimalist design with a "back" feature

EZ Weather uses consistent design throughout the entire interface, where all icons, information, and elements of the same relevance are consistent. Figure 42 shows that all alert settings have the same layout and can be controlled (turned on/off) the exact same way. In addition, EZ Weather employs a minimalist design, where only the necessary information and most frequently used elements are displayed. As shown in

Figure 42, only the most common and critical weather alerts are shown in the alert settings menu. For the "back" feature, EZ Weather enables users to go back one step to access the previous screen (see Figure 42).

12- Ease of manual location change

13- Use of effective search index

14- Flexibility and efficiency of use

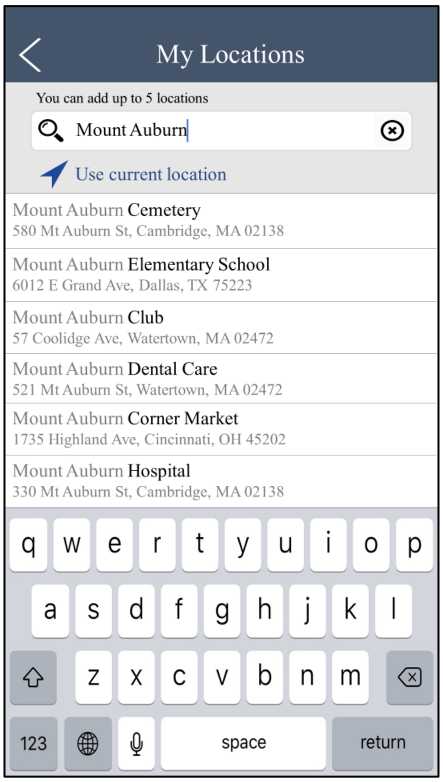


Figure 43. Example of flexible and efficient feature, easy manual location change, and effective search index

In addition to using auto-detection location functionality, EZ Weather enables users to manually change locations by typing location address, name, city & state, or zip code in the app's search bar (see Figure 43). As soon as a user starts typing, an effective search index with auto-suggestions appears to efficiently find and tap a desired location (see Figure 43). Moreover, for flexibility and efficiency purposes, EZ Weather gives experienced and returning users accelerators such as the ability to access recent searched-for locations, instead of having to type again.

15- Visibility of system status

16- User control and freedom

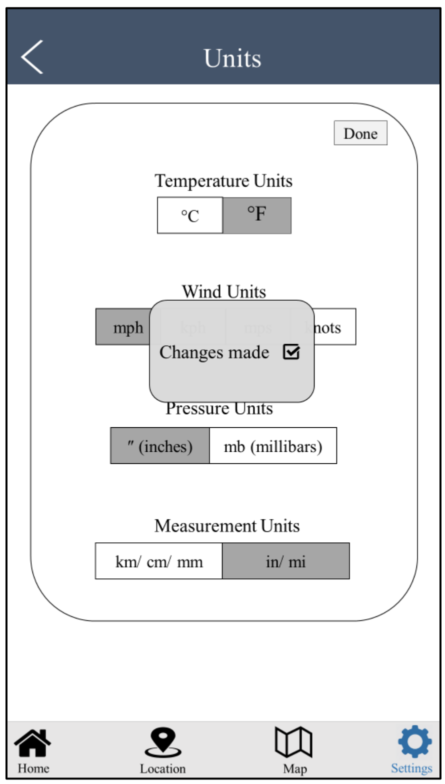
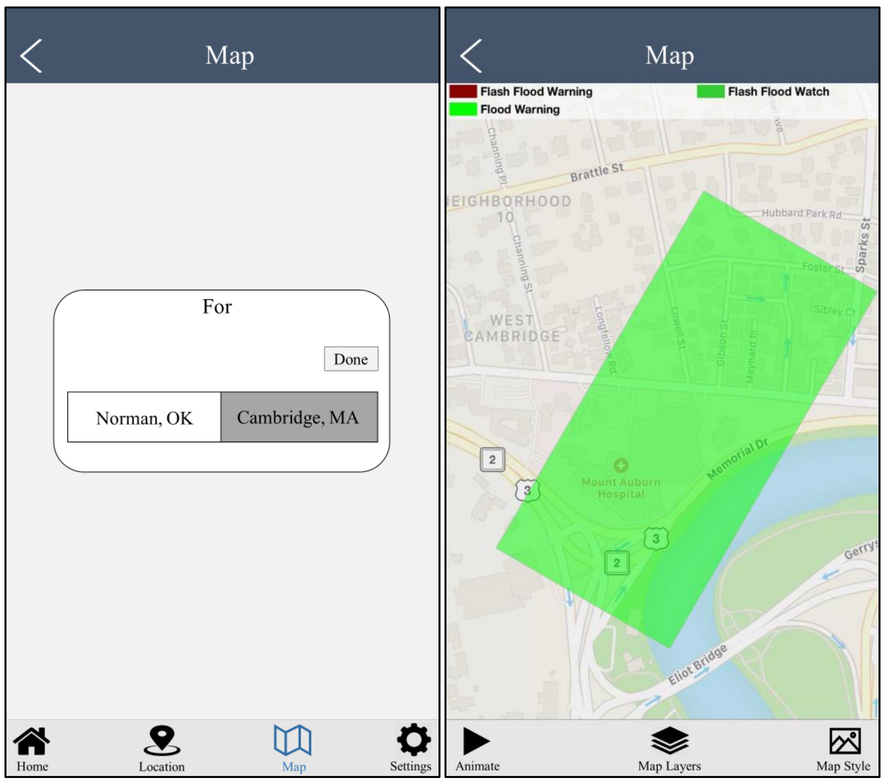


Figure 44. Example of a visible system status and user control & freedom

EZ Weather provides users with visible and concise confirmation messages about any action they perform (e.g. “changes made” after changing units and tapping “done”) as shown in Figure 44. In addition, any icon from the bottom navigation menu highlighted in blue indicates the current used feature such as the settings icon in Figure

44. Furthermore, to enhance users control and freedom, the bottom navigation menu is fixed throughout the entire app so that users can easily and efficiently recover from any unwanted action or mistake.

17- Recognition rather than recall



(a) (b)

Figure 45. Example of visible and intuitive feature

To enhance user recognition, all elements, actions, and information are visible for users and do not require any extensive use of either the short-term or the long-term memory. For example, Figure 45 (a) shows that once a user taps "Map" to visually see an alert on the map and/or control map settings, they are visibly and intuitively given a list of saved locations to choose from; as soon as they choose a desired location, they can see the location’s map along with map setting menus (see Figure 45 (b)). In addition, as previously mentioned, all icons are labeled so that users can intuitively realize the functionality of the displayed elements.

18- Matching between the system and real world

19- High and clear color contrasts

20- Large text font size

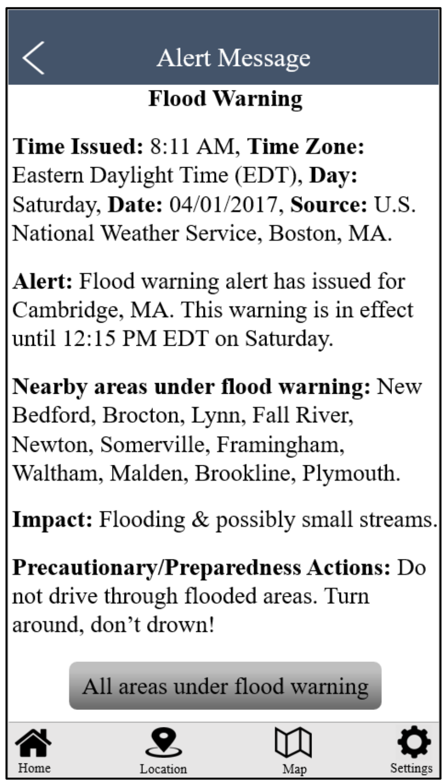


Figure 46. Example of easy language, structured information, clear color contrasts, and large text font sizes

Through the entire EZ Weather app, we used an easy language with everyday words as well as structured information such as in Figure 46. Additionally, we enlarged the text font size according the pre-existing heuristic guidelines (12 point or larger) with even larger font size for the most important and critical information such as the pushed weather alert notifications (see Figure 46). In addition to the large font size, we used high color contrast such as black text on white background, as shown in Figure 46.

In addition to the aforementioned user needs and common heuristics employed in EZ Weather, we used bright screens, limited number of colors, limited number of operational steps, no moving graphics or flashing texts, and no clutter of information. All these used guidelines are believed to substantially result in enhanced user performance and satisfaction with weather alert apps.

To know how EZ Weather addressed users' needs illustrated in Chapter 4, we present a detailed textual and graphical explanation of all enhanced features as follows.

- **Location Search**

As illustrated in Chapter 4, lately, few weather apps have employed an exact location search feature (enabling users to search for and save specific locations), in addition to the traditional search methods (e.g. zip code and/or city), to provide users with precise weather forecasts. However, the efficiency of this feature may have been overlooked. Specifically, these apps restrict users' search for locations by requiring them to navigate and pinpoint the desired location on a map (see example in Figure 3). Hence, the end-users called for an efficient search method, as this feature requires high awareness of the map's geographical area, extensive visual attention, and frequent zoom-in/out within the small smartphone screen.

To address this issue, EZ Weather employs a similar approach as the Google Maps app: typing familiar exact locations with effective auto suggestions in the app's

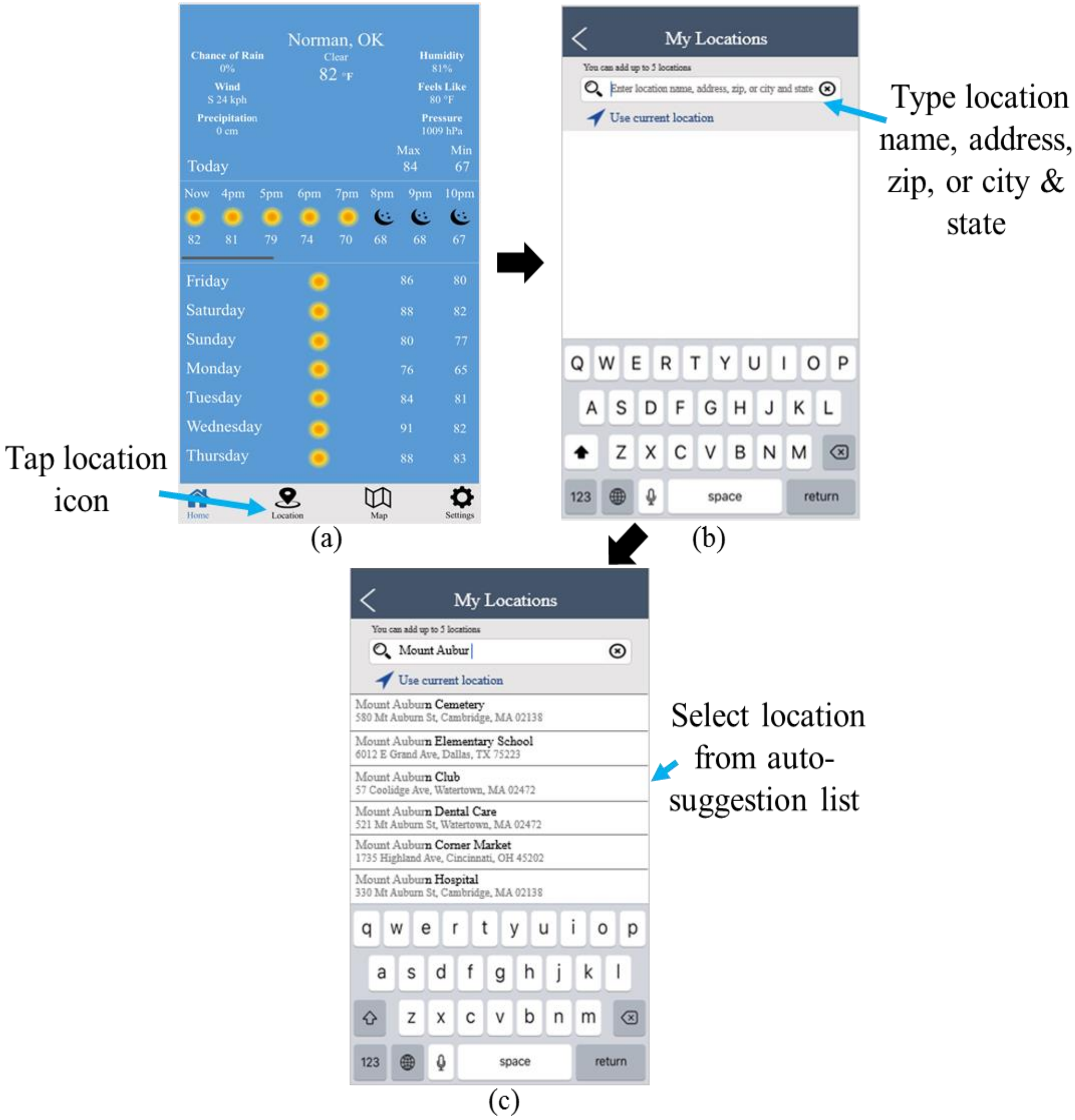


Figure 47. Example of location search process on EZ Weather

search bar. This is believed to substantially enhance the efficiency of this feature. Figure 47 shows the exact process of using this feature on EZ Weather. To add a location, a user has to (a) tap the location icon from the home screen, (b) type the location name (e.g. hospital name/address), address, zip code, or city & state, depending on their needs and preferences. Once a user starts typing, a list of auto-suggested locations appears as in (c) for user selection. The location is saved automatically as soon as it is tapped from the list. It is worth noting that a user can also tap “Use Current Location” from screen (b) to add their current location.

- **Weather Forecasts**

As indicated in Chapter 4, users also voiced concerns about the efficiency of the weather forecasts feature and its priority in weather apps. They stated that many current weather apps require several operational steps on multiple screens to access weather forecasts (see example in Figure 4). Hence, the users indicated a need for limited number of steps to access the weather forecasts feature, as they believed it is more frequently accessed than the other features.

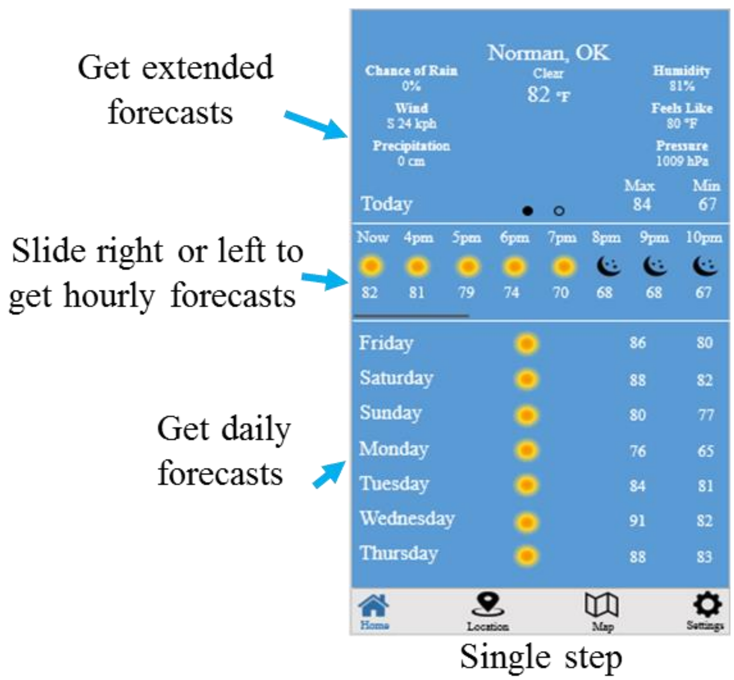


Figure 48. Example of accessing weather process on EZ Weather

EZ Weather considered this feature as among the most important features and placed all weather forecasts (daily, hourly, and extended current forecasts) of each saved location within a single screen. This screen is also used as an introductory or home screen, meaning that a user can access this feature as soon as they open the app. Figure 48 shows the process of accessing the weather forecasts feature on EZ Weather. Once a user is on the desired location, they can see extended weather forecasts (e.g. humidity and feels like) at the top part of the screen and daily forecasts at the bottom part of the screen. Hourly forecasts can be accessed by sliding the middle part right or left to access a 24-hour forecast; the grey line beneath the hourly forecasts area is used as a reference and an indication of the sliding feature. In addition, to view weather forecasts for other saved locations, a user can swipe right or left, with the carousel (little circles) used as a reference; a filled circle indicates a current location.

- **Alert Messages**

Chapter 4 also showed that the focus group participants highly emphasized a need for concise and structured pushed alert messages during severe weather conditions; the alert messages are generated by weather agencies' systems and automatically sent to third parties, including weather apps, where they push them exactly as received to end-users, as visually illustrated in Figure 13. This was due to the fact that the alerts typically contain technical data (e.g. geographical area codes) and cluttered information (see example in Figure 5), which may hinder users' comprehension of the alerts and increase their cognitive load. In addition, the alerts usually include a large amount of information; most of which is unrelated to a user's saved location (e.g. names of and information about all under-alert areas). Even though accessing information about distant or irrelevant locations is critical for some users in specific contexts of use (e.g. when traveling), including them as a main part of the alert may not be to users' benefit.

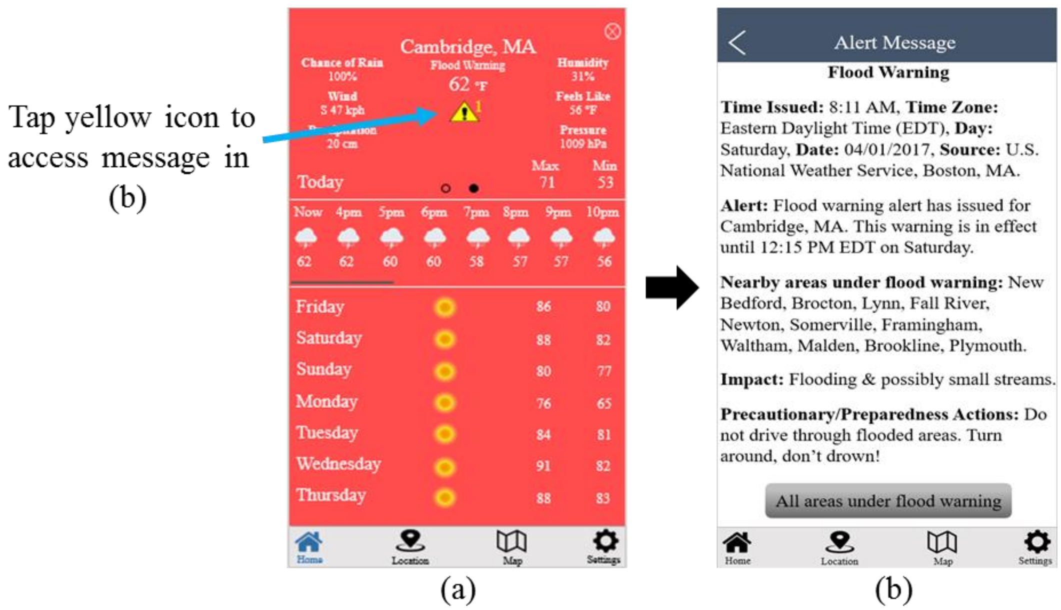


Figure 49. Example of accessing alert messages process on EZ Weather

To address this issue, EZ Weather employed a filtered message content approach by including only the relative and necessary information to the user's saved location on a main alert screen; all other information including distant under-alert areas can be accessed from a secondary menu. Additionally, EZ Weather used simple language of everyday words and hierarchically structured information. This is believed to significantly reduce the user cognitive load and enhance the user comprehension of the messages and reaction to the alert threats. See an example of structured, prioritized, and language-simplified content in Figure 49. In addition, Figure 49 shows the process of accessing the alert messages feature. A user has to tap the yellow alert icon placed on the screen of the affected location, as shown in (a) in order to access the related alert message in (b). Screen (b) only shows the critical information for the user's saved location and nearby areas; information about all other under-alert areas can be accessed on another screen once the bottom option is tapped.

- **Alert Settings**

For the alert settings, weather apps either give users control of alerts and sub-alerts for all weather types (see example in Figure 6) or do not give them control of any alerts and rather automatically push active alerts as notifications. Because of this, the users in the focus groups voiced concerns about the substantial number of weather alert types and sub-alerts, as most of which are rarely needed by average users and/or are not critical. In addition, pushing notifications of any active alert without the end-user's control was perceived as forced interaction. Consequently, the users stated a need for the ability to control only a few relevant alerts.

To avoid the overwhelming number of alerts with which users need to control and interact, as well as to give users freedom to control pushed alert notifications, EZ Weather included only the most critical and common alerts to be controlled (turned on/off) by users. During severe weather conditions, users receive alert notifications for their turned-on time-critical alerts (e.g. tornado warning). Non-time-critical alerts (e.g. wind watch) are not automatically sent to users as notifications; those alerts only appear after tapping a representative symbol on the affected location's screen (see example in Figure 50). Figure 50 also shows the process of accessing and controlling the alert settings on EZ Weather. A user has to (a) tap the settings icon, (b) tap “My Alerts” from the general settings menu in (b), and (c) turn on/off alerts based on their preferences.

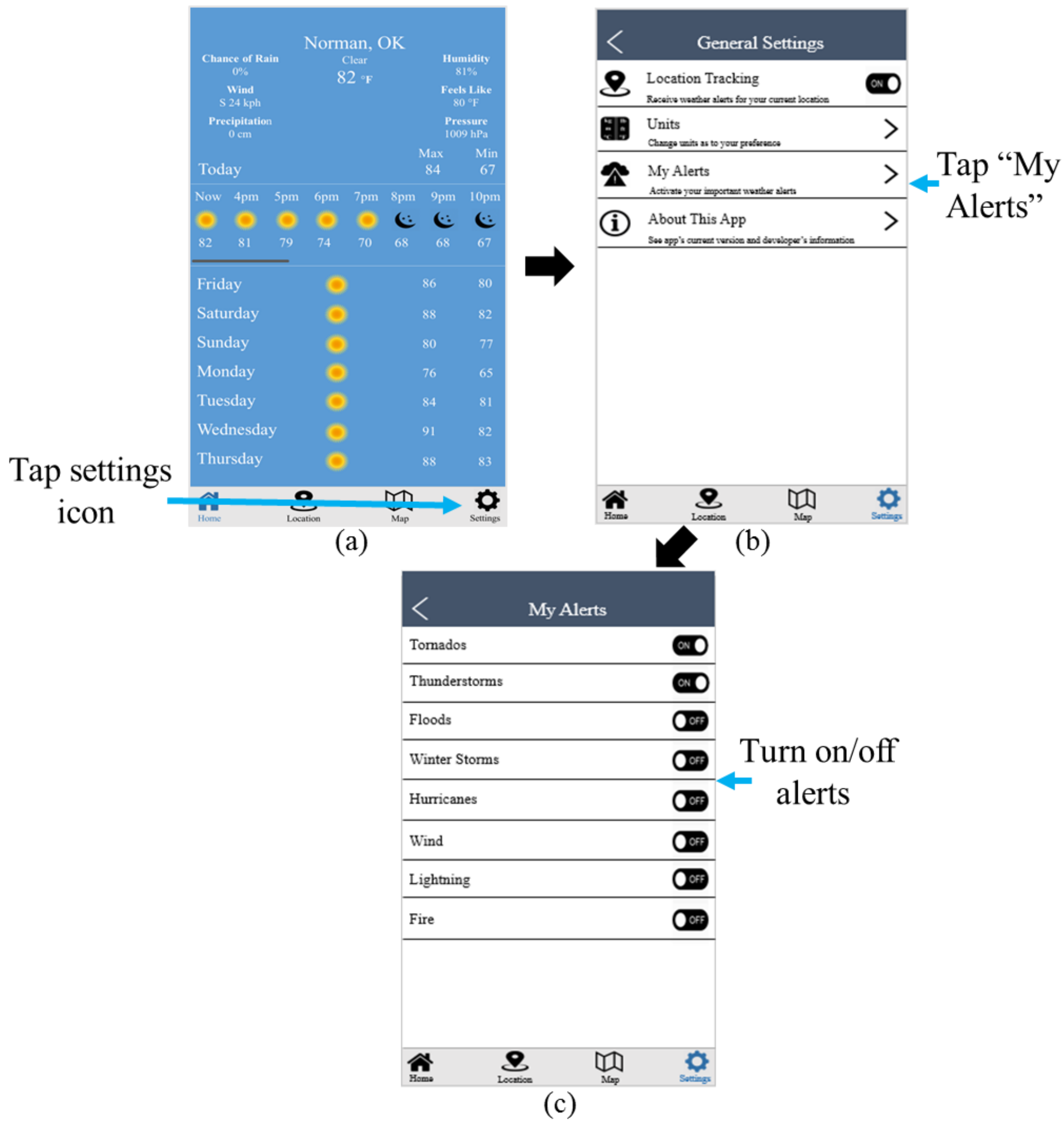


Figure 50. Example of controlling alert settings on EZ Weather

- **Radar Maps**

Chapter 4 also shows that the users called for a radar map feature that is easy to recognize and understand. They shared examples of several popular weather apps that lacked intuitive and/or visible indications on how to access the maps of their respective saved locations or control their settings. Specifically, current weather apps either restrict users to add one location at a time and then access its map or allow them to add multiple locations and do not show indications of how to access their maps. See example of counterintuitive and invisible process to access both saved locations' maps along with their menus in Figure 7.

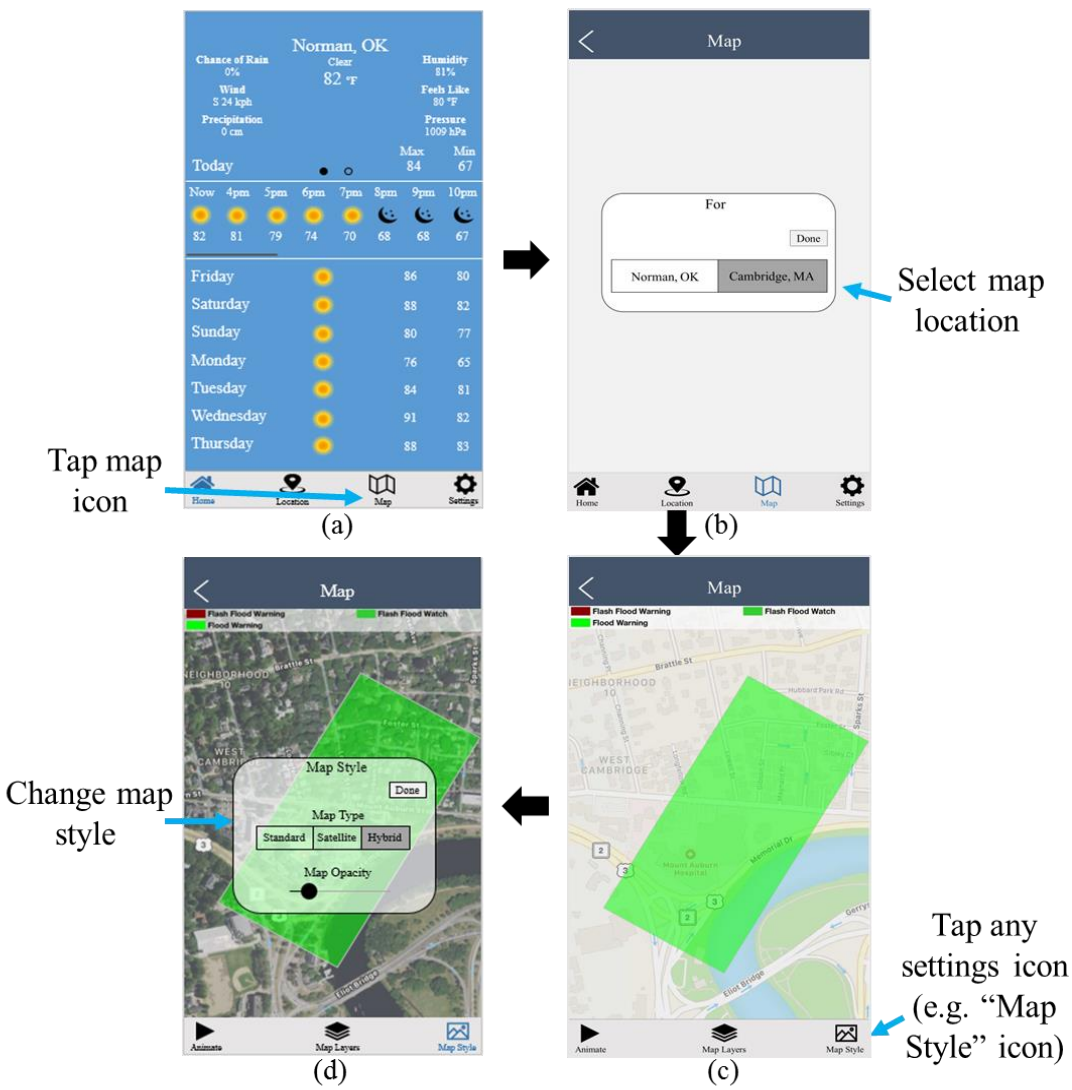


Figure 51. Example of accessing radar map and controlling map settings on EZ Weather

To ease users' recognition and understandability of the radar map feature, EZ Weather displays a list of all saved locations; from which a user can visibly and intuitively select the desired one; the selected location's map along with its settings are then shown. In addition, as EZ Weather labels all icons with representative appropriate text, the sequence of operations is expected to follow a logical order; see example in Figure 51. Figure 51 also shows the process of viewing weather conditions on radars and controlling radar map settings on EZ Weather. To do that, a user has to (a) tap the map icon, (b) choose a desired location from the list of saved locations to see the radar, (c) tap any labeled icon at the bottom of navigation menu to control the corresponding map settings, and (d) access the corresponding settings and change as to their preferences.

5.3 Chapter Summary

This chapter has presented the developed smart phone prototype weather app, named "EZ Weather", which is the second phase in the UCD process. In the developed prototype app, we addressed all user needs voiced in the focus group interviews and explained in Chapter 4. In addition, we used the knowledge attained from the preliminary findings in Chapter 3 in the interface design of the UCD prototype app. Moreover, we took into consideration all general usability heuristics introduced by Nielsen (1995a), except "displaying error messages". This was due to the limitation of prototype interfaces as error messages can only be displayed on running systems. In addition to error messages, we did not provide a "help and documentation" feature as Nielsen (1995a) believed that a usable system should not include this heuristic unless the interface includes complex features in nature. We believe that our interface does not contain any complex features. Furthermore, we accounted for all older users' limitations listed in Chapter 4 and addressed them in EZ Weather following the heuristics shown in the same chapter. Moreover, we employed the specific iOS smart phone app design heuristics introduced by Kaufman, (2016), Gove, (2016), and Griffiths, (2015).

By reaching this stage, we believe that a validation usability study to examine whether the developed app has practically addressed end-user needs, is required.

Chapter 6. Usability Evaluation of Older and Younger Users' Experiences with the Smart Phone Prototype Weather Application (Phase 3)

Author's Note: The content in this chapter will be submitted soon (Status: final edits) as a journal paper to the Journal of Human Factors and Ergonomic Society (HFES). The author of this dissertation wrote this paper in collaboration with his advisor, Dr. Ziho Kang.

6.1 Introduction

This chapter presents the third phase in the UCD process, which is testing the usability of the developed UCD prototype smart phone weather app (EZ Weather). After identifying users' needs, limitations, and characteristics in the first UCD phase (see Chapter 4) and developing a prototype app based on users' feedback in the second UCD phase (see Chapter 5), we aim in the third phase to determine whether the developed UCD app greatly enhances users' experiences with weather apps.

To perform a valid and reliable usability evaluation, we benchmark EZ Weather with the widely used smart phone weather app, Weather Radio. In addition, similar to the approach followed in Chapter 3, we target first-time users for evaluation. However, in this study we evaluate the usability of the UCD app with two age groups of first-time users: younger and older users. It is worth reminding that we considered needs, limitations, capabilities, and heuristic guidelines for both younger and older users in the previous two chapters; hence, one of our goals in this study is to determine to what extent the UCD app is user-friendly for all end-users, regardless of age.

All the enhanced five features in the UCD app are evaluated in this study: (1) location search (i.e. typing familiar locations with effective auto-suggestions), (2) weather forecasts (i.e. all-inclusive weather forecasts within one screen), (3) alert messages (i.e. use of structured, prioritized, and language-simplified alert messages content), (4) map settings (i.e. use of visible and intuitive map menus), and (5) alert settings (i.e. use of minimalist alert settings), and are compared to their analogous features on Weather Radio. For evaluation, we use the ISO 9241-11 (1997) model: effectiveness (measured by task completion rate, number of errors, severity ratings of errors, and causes of errors), efficiency of use (measured by task completion time), and user satisfaction (measured by post-task satisfaction survey and post-test satisfaction survey). Finally, we investigate the correlations among the evaluation metrics in order to determine if there are specific relationship trends that can be concluded.

6.2 Method

Participants

Eighty regular iOS smartphone weather app users were recruited for the experiment. Both younger (*M* = 25.9 years, *SD* = 4.8, and *Range* = 18 – 35 years) and older (*M* = 57.4 years, *SD* = 4.3, and *Range* = 50 – 66 years) users participated in the experiment. The users were randomly assigned to perform tasks on the two tested apps (Weather Radio & EZ Weather). Each app was used by 40 users (20 younger users and 20 older users). To perform a reliable and valid experiment, the experimenters made sure that all participants were iPhone users with at least six months of use, first-time users of the tested apps, and active users of other smart phone weather apps. Recruitment was based on personal communication, university’s mass email, and flyers hung on the doors of various public buildings.

Apparatus

Both Weather Radio (version 3.0.5) and EZ Weather (version 1.0.0) were installed and operated on an iPhone 6. A high-fidelity simulation using a powerful interaction design system, called InVision, was used to show the recorded alert message of Weather Radio at any time during the experiment's time frame. A Nikon L340 camera was used to video record users' interactions and particularly measure the time to complete the given tasks, as well as to count and categorize errors. The demographic, post-task, and post-test surveys were printed out on paper.

Scenario & Tasks

Figures 3 to 7 show examples of the steps needed to complete all tasks on Weather Radio and Figures 45 to 49 show the steps required for all tasks on EZ Weather. The participants were given a scenario where their grandmother was an in-patient at "*Mount Auburn hospital in Cambridge, MA"* and that they became aware of a flood warning alert issued for Cambridge area. To know the alert's risk-level for the grandmother's exact location and access all related information, the participants needed to search and add her exact location, access relevant weather forecasts & alert messages, see the alert on map with specific map settings, and adjust alert settings to receive relevant alert notifications. The tasks given to the participants were as follows:

• Task 1 (Location Search). *Add "Mount Auburn hospital, Cambridge, MA" to your saved locations*.

• Task 2 (Weather Forecasts). *Access the location (Cambridge, MA), find, and verbally state current wind and feels like forecasts*.

• Task 3 (Alert Messages). *Access the alert message associated with Cambridge, MA location and verbally answer the following questions: (1) What impact might the flooding cause? (2) What action needs to be taken? (3) What time does the flood warning expire?*

• Task 4 (Map Settings). *Adjust map settings to be able to see the alert area on a hybrid map view.*

• Task 5 (Alert Settings). *Adjust alert settings by turning on the alerts of "Severe Thunderstorms" and "Floods".*

Procedure

First, each participant signed a participation consent form and filled out a demographic survey. Then, the participants were given a brief description of the study's objective. All participants neither received training nor were given pre-self-practice chances with the tested apps; they were given the scenario and tasks instructions and asked to begin the experiment once they informed readiness. All five tasks were administered on both apps. However, as Weather Radio was a running app and alert messages would not be shown except during active alerts, Weather Radio's alert messages task was presented to participants on a smartphone interface designed through InVision; the interface layout, color, and message content were identical to those on the actual Weather Radio interface. The alert message displayed to the participants was a flood warning alert pushed to weather app users on April 1st, 2017 in Cambridge, MA. After performing each task, the participants completed a post-task survey to express their instant feelings about each task/feature. Finally, the participants completed a post-test survey to rate their satisfaction levels with the tested apps.

Experimental Design & Variables

The experiment followed a 2*2 Between-Subject design. This design was used to avoid any learning effect during the experiment. The first independent variable was app used with two levels: Weather Radio and EZ Weather and the second variable was age group with two levels: younger and older users. The dependent variables were effectiveness (measures: task completion rate and errors-related metrics), efficiency (measure: task completion time), and user satisfaction (measures: post-task and post-test satisfaction surveys). The error metrics were causes (usability problems) of the made errors, their frequency of occurrence, proportions of users who made them, as well as severity ratings of the error causes. The severity ratings of the error causes used in this study are based on a rating scale proposed by Nielsen (1995c) as shown in Table 8. The evaluation of errors based on their severity ratings was performed by two independent usability practitioners.

Table 8. Nielsen's severity rating scale of the usability problems

0	I don't agree that this is a usability problem at all
1	Cosmetic problem only: need not be fixed unless extra time is available on project
2	Minor usability problem: fixing this should be given low priority
3	Major usability problem: important to fix, so should be given high priority
4	Usability catastrophe: imperative to fix this before product can be released

The post-task survey was a Single Ease Question (SEQ) with 7-point Likert rating scale: **"this task was**

Very Difficult 1 2 3 4 5 6 7 Very Easy".

This question was used as it was found to be as effective as other complicated metrics (i.e. Usability Magnitude Estimation (UME) and Subjective Mental Effort (SMEQ) questionnaires) of task-difficulty (Sauro & Dumas, 2009). The post-test survey adapted in this study was the Questionnaire for User Interface Satisfaction (QUIS) (Chin et al., 1988). The QUIS survey items are shown along with their relevant results in the results section. We adapted this questionnaire in this study because of its comprehensive and specific approach to analyze several important and critical areas in interface designs such as screen design, terminology, the user's learning experience, system capabilities, and multimedia.

The control variable was user experience, which all users in both age groups were required to have an experience with iOS smartphone devices and weather apps (excluding the tested apps) for at least six months.

Data Analysis

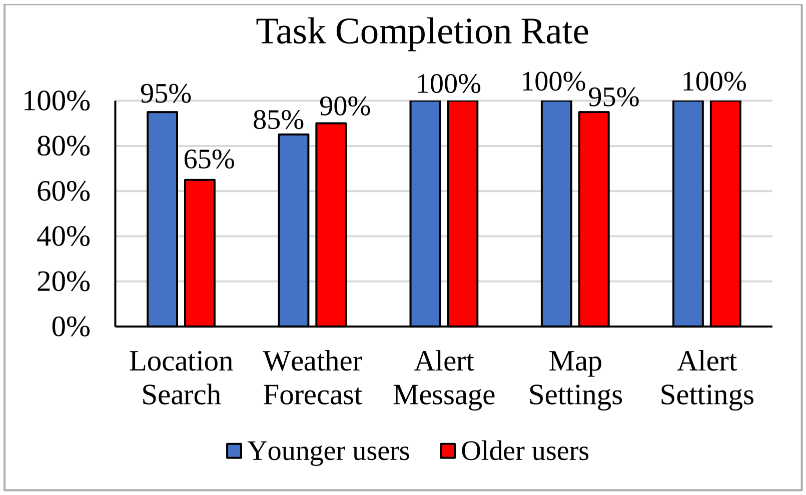
All collected data from video recordings and survey sheets were analyzed using SPSS Version 23. A two-way Between-Subject analysis of variance (ANOVA) was conducted to determine the effect of app used and age group on the task completion time for all experiment's tasks. Due to the data violation to parametric test assumptions, a Mann-Whitney U test was used to examine the differences magnitude among app used and age group variables, in terms of both number of errors and post-task satisfaction survey metrics. To analyze tasks completion rate, proportions of successfully completed tasks by users were used. For the post-test satisfaction survey, the mean and standard errors for each survey item were calculated. Finally, the standard Pearson correlation test (*r*) was performed to determine the association levels among the used usability measures in this study.

6.3 Results

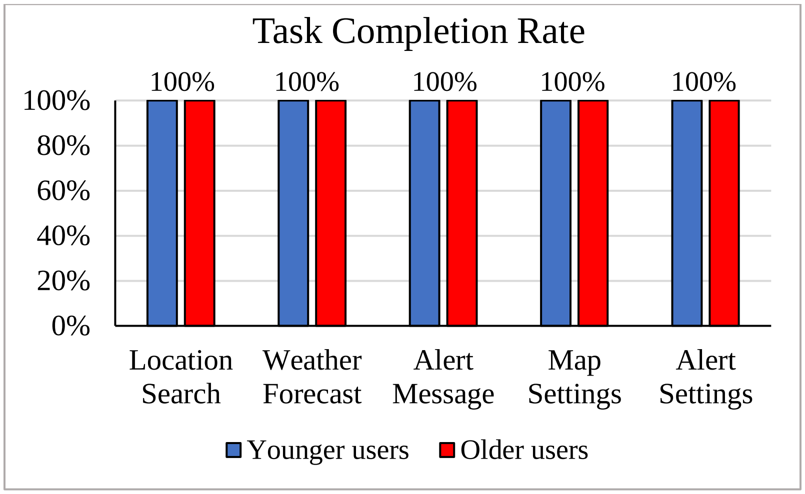
Effectiveness

- **Task completion rate**

The proportions of successfully completed tasks for both age groups on both apps were calculated and reported as shown in Figure 52 (a & b). It was shown that all users in both age groups were able to successfully complete the given tasks on EZ Weather. In contrast, several users failed to successfully complete three tasks on Weather Radio (location search, weather forecasts, and map settings), with a larger failure rate among older users on the location search task.



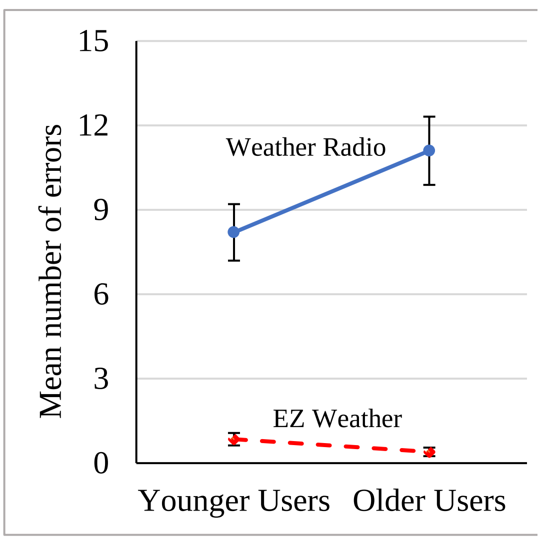
(a) Weather Radio



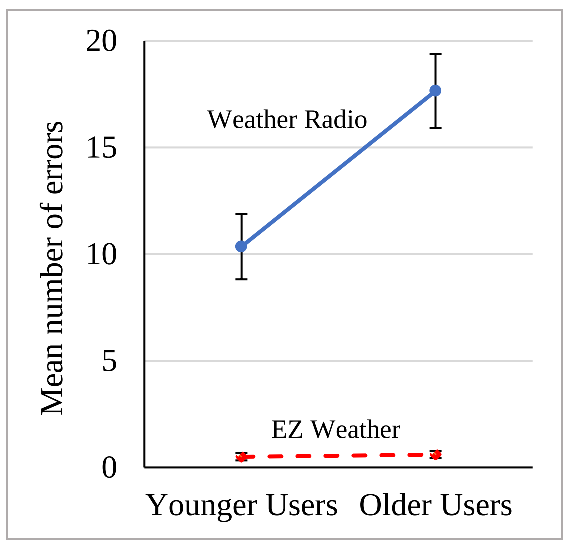
(b) EZ Weather

Figure 52. Proportions of successful task completion for both age groups on (a) Weather Radio and (b) EZ Weather

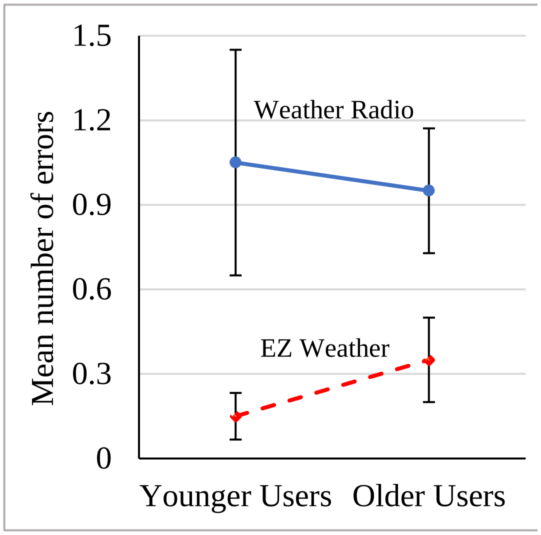
- **Error**



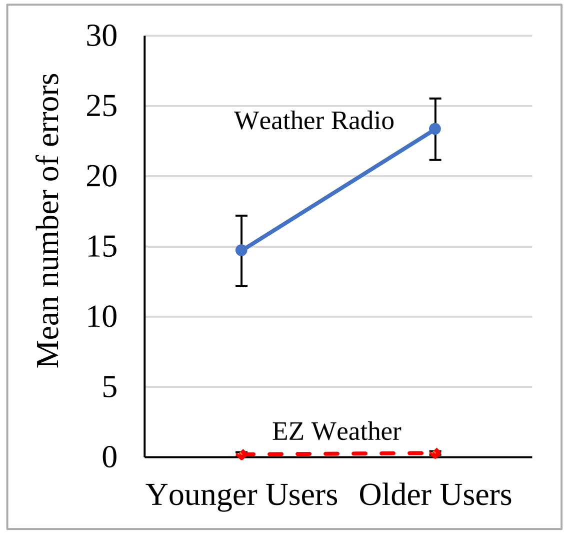
(a) Location search



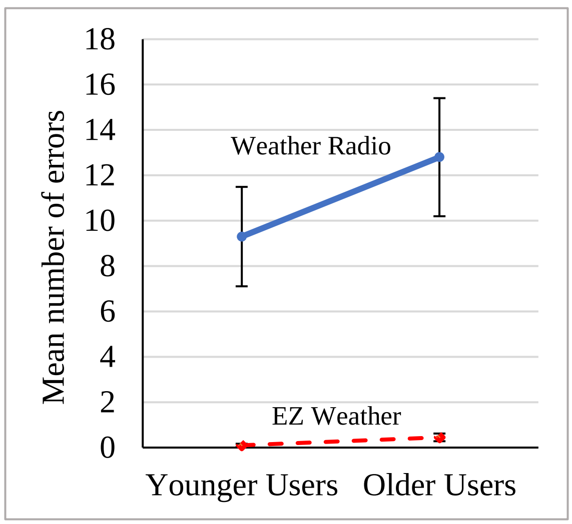
(b) Weather forecasts



(c) Alert messages



(d) Map settings



(e) Alert settings

Figure 53. Mean number of errors for both age groups on all tasks (a to e)

Figure 53 (a-e) shows that for all tasks, both younger and older users made noticeably fewer errors on EZ Weather than they did on Weather Radio. The results also indicate that older users made more errors on both apps than did younger users on all tasks, except on the alert messages task of Weather Radio and the location search task of EZ Weather.To test the differences among the levels of the independent variables, a Mann-Whitney U test was performed and indicated an existence of a significant difference between the two apps for all tasks, in terms of the number of errors (see Table 9).

The results also showed that no significant error difference was observed between the two age groups for all tasks, except for the weather forecasts and map settings tasks. Due to limitations with the Mann-Whitney test, it was not possible for us to calculate the interaction effect.

Table 9. Mann-Whitney test summary for number of errors

Task	Source	Z-score	U-test	P-value
Location Search	Age group	-1.22	677.50	.223
	App used	-7.86	9	<.001
Weather Forecast	Age group	-2.14	582	.033
	App used	-6.64	123	<.001
Alert Message	Age group	-.56	743	.575
	App used	-7.68	20	<.001
Map Settings	Age group	-2.42	655.5	.015
	App used	-7.51	34.5	<.001
Alert Settings	Age group	-.91	710.5	.364
	App used	-8.01	10	<.001

In terms of the other error-related metrics, almost all users in both age groups made errors, with a different frequency, due to the same usability problems for each task on Weather Radio, except for the alert messages task (see Table 10). Nearly half of the users made errors because of the alert message usability issues.

In contrast, a smaller number of users made errors on EZ Weather’s tasks with substantially less frequency of occurrence and average severity ratings compared to those on Weather Radio. The errors made on EZ Weather were mostly caused by slip actions (e.g. typos and accidently tapping adjacent function icons). See Table 11 for details.

Table 10. Causes of errors, frequency of issue, (proportions of users who made errors), and average severity ratings on Weather Radio

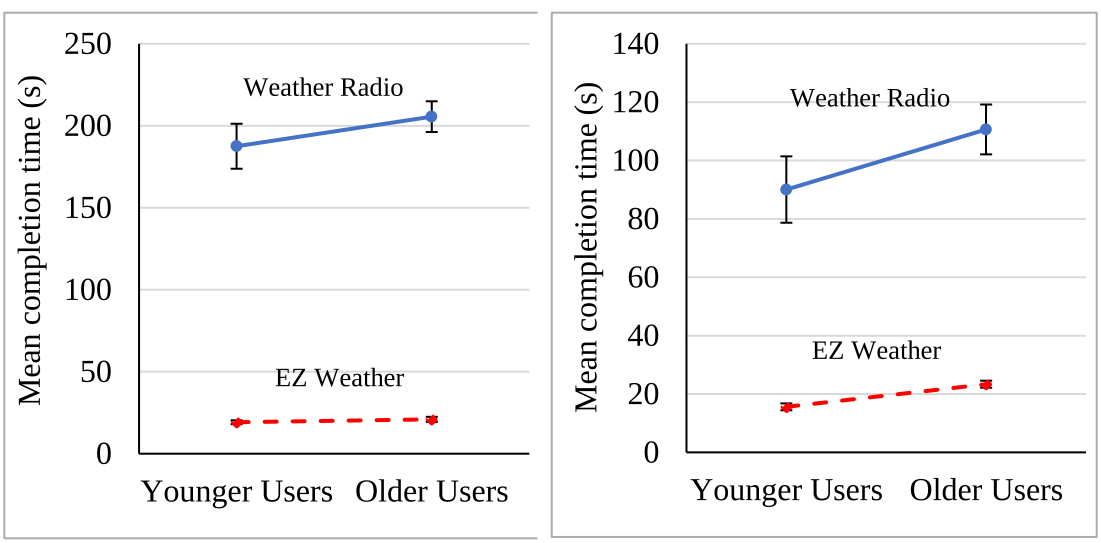
Feature	Cause of errors (Usability problem)	Frequency of issue		Ave. Severity Rating
		Younger Users	Older Users	
Location Search	Users had trouble finding location, as well as moving the pin on the map.	178 (100%)	253 (85%)	3.5
Weather Forecast	Users couldn't easily locate weather forecasts. The area leading to forecasts, if clicked, didn't seem to be clickable.	231 (100%)	366 (100%)	3.5
Alert Message	Users couldn't easily access the necessary information of the time-critical weather alert message because of cluttered & unstructured information and poor use of language.	21 (40%)	17 (45%)	4
Map Settings	Users had difficulty beginning the task because of the counter-intuitive steps and the invisible map settings menu.	166 (80%)	287 (100%)	3
	Users didn't understand the functionality of the map settings icons as the icons were neither labeled nor standardized across smart phone apps.	126 (90%)	194 (100%)	2.5
Alert Settings	Users did not understand the functionality of the home screen icons as the icons were neither labeled nor standardized across smartphone apps. Also, the substantial number of alerts & sub-alerts seemed to confuse users about the required options for the tasks.	186 (90%)	256 (95%)	3.5

Table 11. Causes of errors, frequency of issues, (proportions of users who made errors), and average severity ratings on EZ Weather

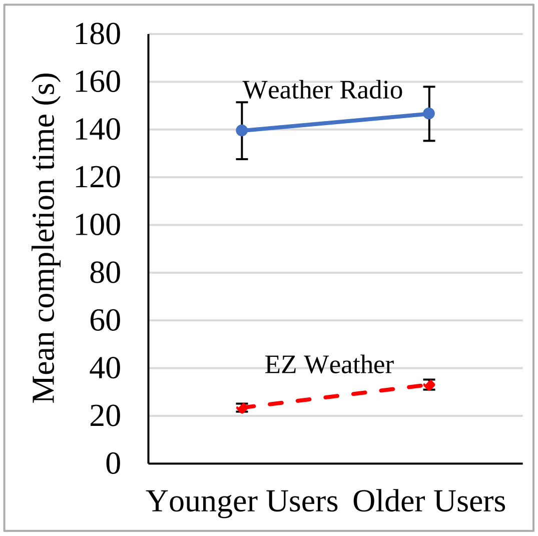
Feature	Cause of errors (Usability problem)	Frequency of issue		Ave. Severity Rating
		Younger Users	Older Users	
Location Search	Users made typing errors when typing location name.	17 (55%)	9 (30%)	0
Weather Forecast	Users couldn't easily figure out that accessing weather forecasts of different locations was through swiping the screen right or left.	10 (35%)	12 (45%)	1.5
Alert Message	Users didn't expect the alert message icon to be clickable and/or required to access the message, when clicked.	3 (15%)	7 (25%)	1
Map Settings	Users mistakenly tapped adjacent icons of unrelated functions.	4 (10%)	6 (25%)	0
Alert Settings	Users mistakenly tapped adjacent icons of unrelated functions.	2 (10%)	9 (35%)	0

Efficiency

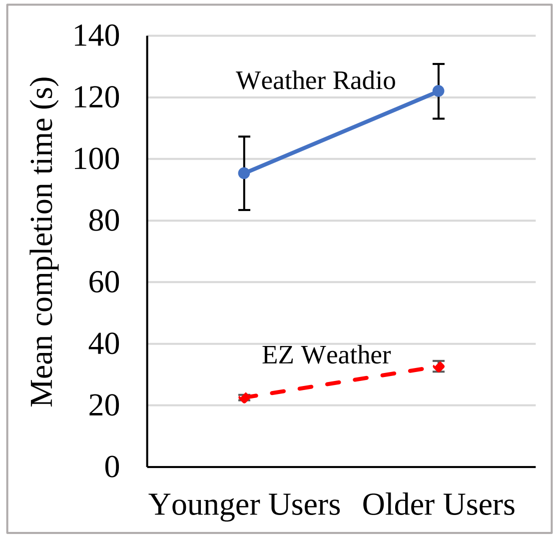
- **Task completion time**



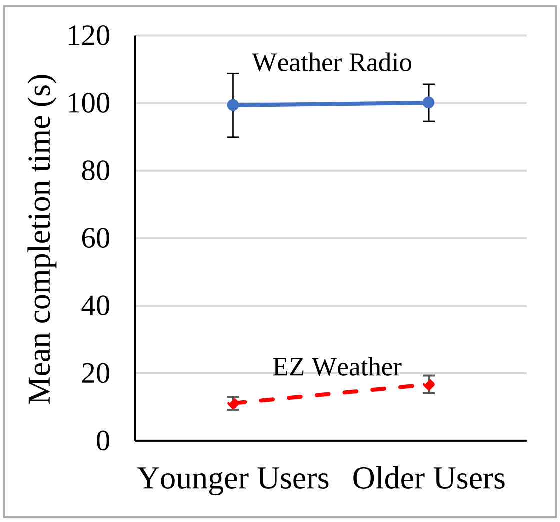
(a) Location search (b) Weather forecasts



(c) Alert messages



(d) Map settings



(e) Alert settings

Figure 54. Mean task completion time spent in completion of all tasks (a to e)

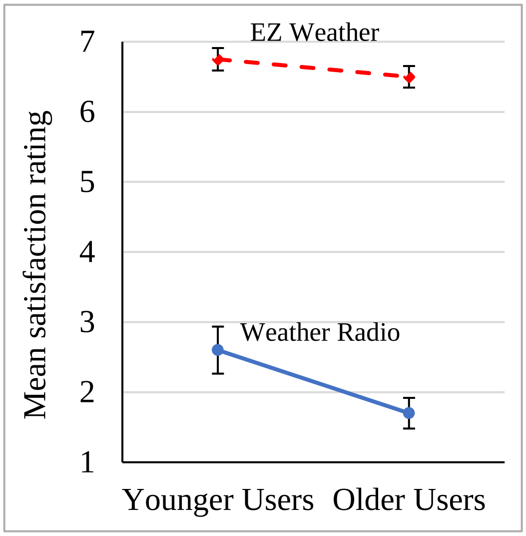
Figure 54 (a-e) shows that, on average, both younger and older users needed substantially less time to complete each of EZ Weather's tasks than did on Weather Radio. The time was relatively similar for both age groups on all tasks for both apps.

To determine the difference among the levels of the independent variables, a Two-Way ANOVA test (see Table 12) was performed and showed that there was no significant time difference between younger and older users on both weather apps for all tasks, except for the map settings task. However, the time to complete all tasks for both age groups was significantly different across the tested weather apps. The results also revealed that there was no interaction between age group and app used on all tasks, meaning that there was insufficient evidence to reject the interaction effect null hypothesis.

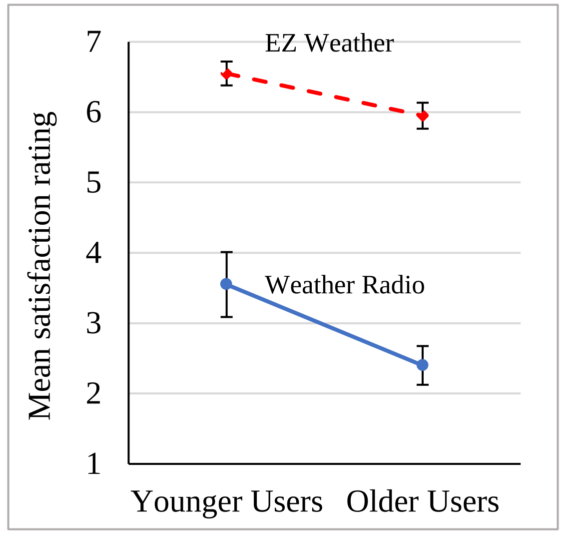
Table 12. Two-way ANOVA summary for task completion time

Task	Source	SS	df	MS	F	P-value
Location Search	User group	1958.19	1	1958.19	1.401	.240
	App Used	622646.43	1	622646.43	445.554	<.001
	User * App	1326.80	1	1326.80	.949	.333
	Within (error)	106207.30	76	1397.46		
	total	1670393.92	80			
Weather Forecast	User group	4019.94	1	4019.94	3.916	.051
	App Used	130777.57	1	130777.57	127.384	<.001
	User * App	826.21	1	826.21	.805	.373
	Within (error)	78024.63	76	1026.64		
	Total	500753.04	80			
Alert Message	User group	1402.81	1	1402.81	1.010	.318
	App Used	263466.01	1	263466.01	189.608	<.001
	User * App	32.51	1	32.51	.023	.879
	Within (error)	105604.55	76	1389.53		
	Total	957551	80			
Map Settings	User group	6755.37	1	6755.37	6.012	.017
	App Used	131312.30	1	131312.30	116.856	<.001
	User * App	1354.16	1	1354.16	1.205	.276
	Within (error)	85401.64	76	1123.71		
	Total	596259.96	80			
Alert Settings	User group	201.61	1	201.61	.332	.566
	App Used	147318.61	1	147318.61	242.835	<.001
	User * App	117.61	1	117.61	.194	.661
	Within (error)	46106.35	76	606.66		
	Total	451957	80			

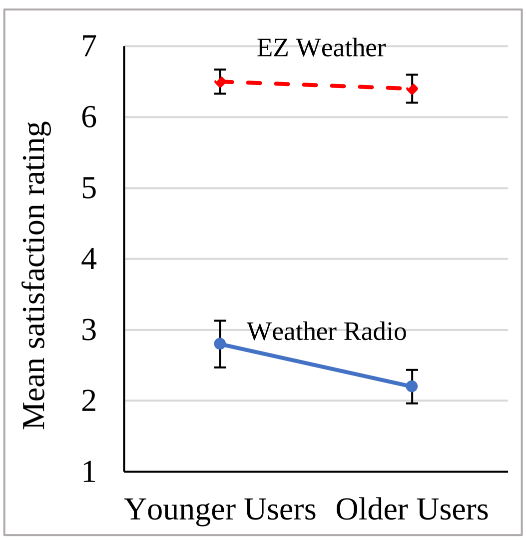
Post-task satisfaction ratings



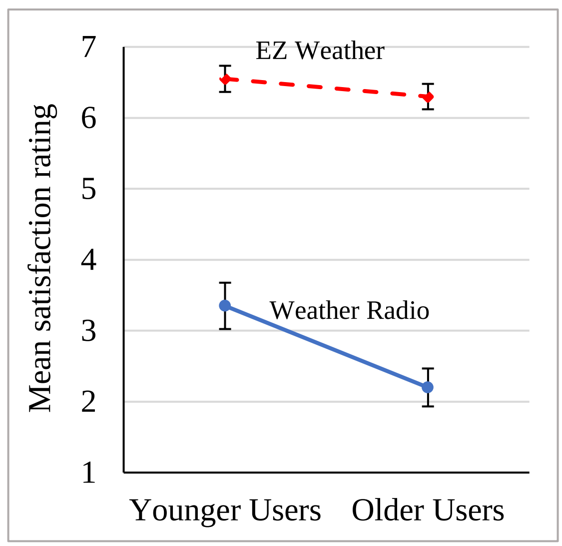
(a) Location search



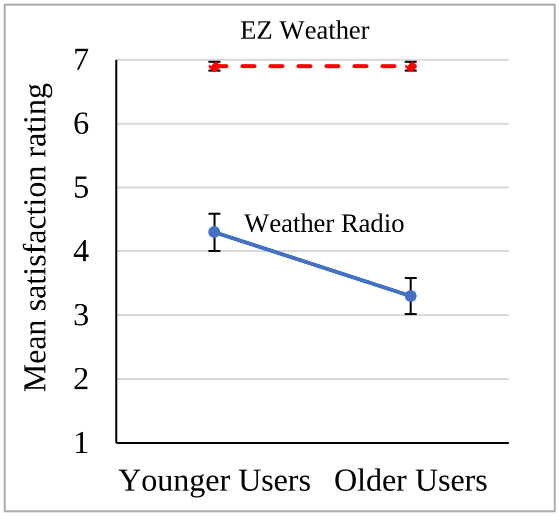
(b) Weather forecasts



(c) Alert messages



(d) Map settings



(e) Alert settings

Figure 55. Mean post-task satisfaction ratings of SEQ survey for all tasks (a to e)

Figure 55 (a-e) shows that both younger and older users believed that EZ Weather tasks were substantially easier than Weather Radio tasks. Younger users rated the ease of tasks on both apps higher than older users, although not by a large margin.

To examine the nature of difference among the levels of the independent variables, in terms of the post-task satisfaction ratings, a Mann-Whitney U test was performed and showed that there were statistically significant task satisfaction differences between the two apps for all tasks, given that the *p-values* for the “app used” variable on all tasks were $< .05$ (see Table 13). On the other hand, Table 13 shows that both younger and older users had similar satisfaction levels for all tasks, given that the *p-values* for the “app used” variable on all tasks were $> .05$. The interaction effect was not calculated due to limitations with the Mann-Whitney test. However, the graphs in Figure 55 indicate that no interaction would be possible between the two variables.

Table 13. Mann-Whitney test summary for post-task satisfaction ratings

Task	Source	Z-score	U-test	P-value
Location Search	Age group	-.298	769.5	.766
	App used	-7.76	5	<.001
Weather Forecast	Age group	-1.32	665	.187
	App used	-7.58	25	<.001
Alert Message	Age group	-.842	727.5	.401
	App used	-2.68	569	.007
Map Settings	Age group	-1.33	667	.184
	App used	-7.97	1.5	<.001
Alert Settings	Age group	-1.32	668	.185
	App used	-7.09	93.5	<.001

Post-test satisfaction ratings

The post-test satisfaction survey (QUIS) tested users' satisfaction levels, in terms of 5 categories: overall reaction (see Table 14), screen, terminology & system information (see Table 15), learning (see Table 16), and system capabilities & multimedia (see Table 17). Overall, the results revealed that both younger and older users had relatively similar feelings towards the interface design specifications of the two tested apps, with higher satisfaction levels by younger users. The results also show that all users were extremely satisfied with EZ Weather, whereas they were mostly disappointed with Weather Radio.

For more details, each category's questions are shown in separate tables and followed by its results. The survey questions and results are as follows.

Table 14. QUIS overall reaction questions

Question		1 2 3 4 5 6 7 8 9		
Overall Reaction	1.	Terrible	Wonderful	NA
	2.	Frustrating	Satisfying	NA
	3.	Dull	Stimulating	NA
	4.	Difficult	Easy	NA
	5.	Adequate Power	Adequate power	NA
	6.	Rigid	Flexible	NA

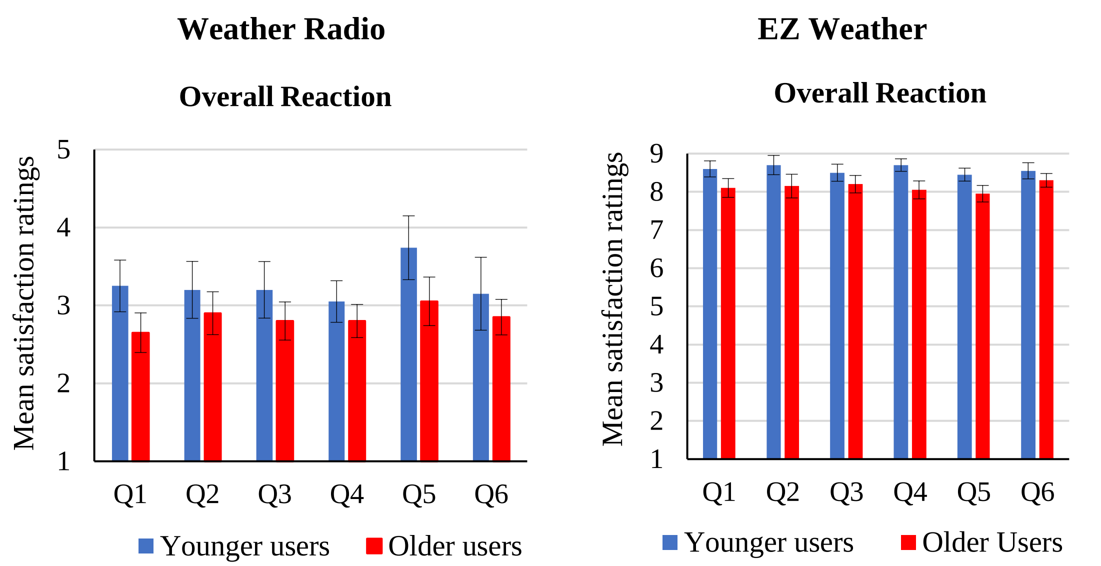


Figure 56. Users' mean satisfaction ratings for the overall reaction category on Weather Radio and EZ Weather

The overall reaction results (see Figure 56) showed that both younger and older users believed that the EZ Weather's interface was extremely wonderful, satisfying, simulating, easy to use, flexible and powerful. On the other hand, both age groups had less than average satisfaction levels regarding the overall reaction questions.

Table 15. QUIS screen questions

	Question	1 2 3 4 5 6 7 8 9		
Screen	7. Characteristics on the computer screen	Hard to read	Easy to read	NA
	8. Image of characters	Fuzzy	Sharp	NA
	9. Character shapes (fonts)	Barely legible	Very legible	NA
	10. Screen layouts were helpful	Never	Always	NA
	11. Amount of information that can be displayed on screen	Inadequate	Adequate	NA
	12. Arrangement of information on screen	Illogical	Logical	NA
	13. Sequence of screens	Confusing	Clear	NA
	14. Next screen in a sequence	Unpredictable	Predictable	NA
	15. Going back to the previous screen	Impossible	Easy	NA
	16. Progression of work-related tasks	Confusing	Clearly marked	NA

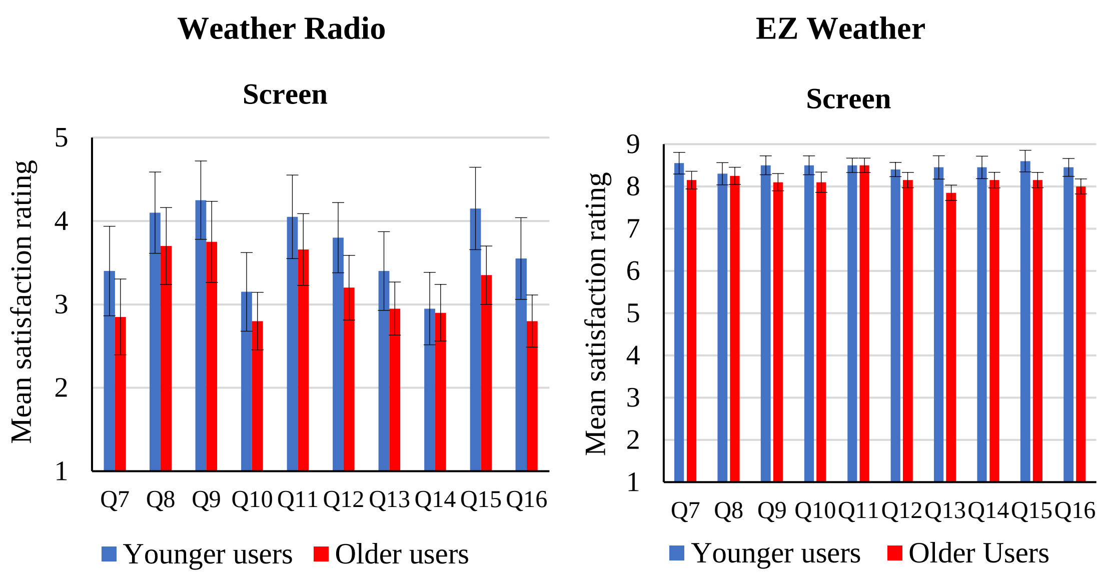


Figure 57. Users' mean satisfaction ratings for the screen category on Weather Radio and EZ Weather

As shown in Figure 57, users reported that the screen of EZ Weather was well designed, as the characters were easy to read, images were very sharp, fonts were very legible, amount of presented information was adequate & arranged in a logical order, sequence of screens and progression of tasks were very easy, and next screen was always predictable. Regarding the amount of information presented on screen, one user stated: *"I really liked this app [EZ Weather], as it shows me exactly what I need in a very concise manner."* In contrast, users appeared to be dissatisfied with Weather Radio's screen design. They commented that it had too much information (especially on the time-critical alert messages), hard to read characteristics, fuzzy images, confusing layouts & sequence of screens, and illogical information arrangement.

Table 16. QUIS terminology & system information questions

	Question	1 2 3 4 5 6 7 8 9		
Terminology & System Information	17. Use of terminology throughout system	Inconsistent	Consistent	NA
	18. Work related terminology	Inconsistent	Consistent	NA
	19. Messages which appear on-screen	Inconsistent	Consistent	NA
	20. Position of instructions on the screen	Inconsistent	Consistent	NA
	21. Messages which appear on-screen	Confusing	Clear	NA
	22. Instructions for commands or functions	Confusing	Clear	NA
	23. Performing an operation leads to a predictable result	Never	Always	NA
	24. Length of delay between operations	Unacceptable	Acceptable	NA

Figure 58 clearly shows that both age groups were very satisfied with the intuitiveness and consistency of the messages positions and terminology used in EZ Weather. They also found that performing an operation would almost always lead to a predictable result on EZ Weather with a very acceptable length of delay between

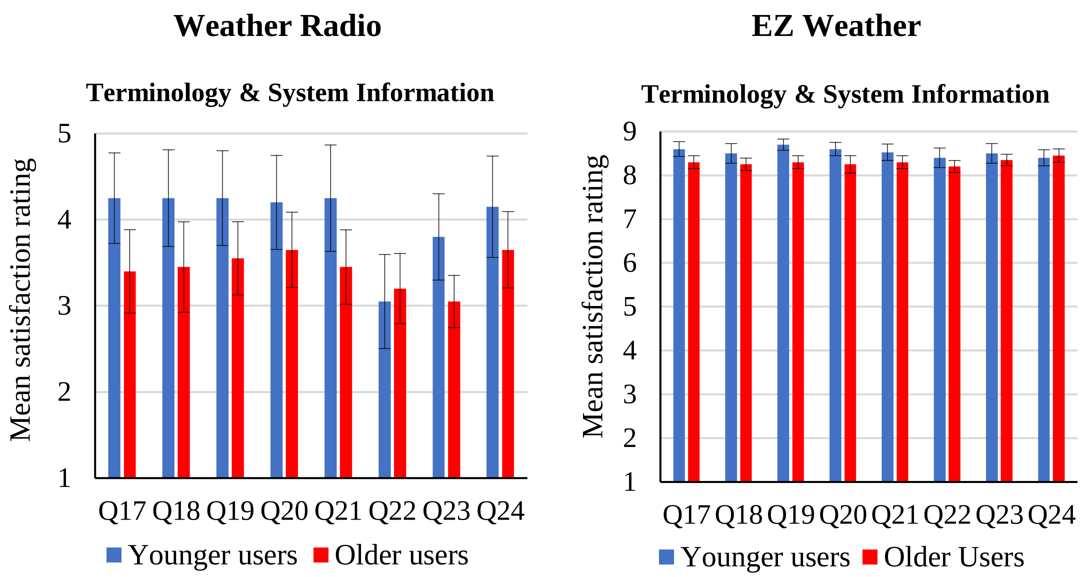


Figure 58. Users' mean satisfaction ratings for the terminology & system information category on Weather Radio and EZ Weather

operations. In contrast, many users reported that they were not sure what to expect when performing several operations on Weather Radio. The use of terminology and positions of messages & instructions were somewhat inconsistent and confusing on Weather Radio. In addition, they believed that the length of delay between operations is relatively unacceptable, compared to that on EZ Weather.

Table 17. QUIS learning questions

	Question	1 2 3 4 5 6 7 8 9		
Learning	25. Learning to operate the system	Difficult	Easy	NA
	26. Getting started	Difficult	Easy	NA
	27. Time to learn to use the system	Slow	Fast	NA
	28. Exploration of features by trial and error	Discouraging	Encouraging	NA
	29. Exploration of features	Risky	Safe	NA
	30. Discovering new features	Difficult	Easy	NA
	31. Remembering names and use of commands	Difficult	Easy	NA
	32. Remembering specific rules about entering commands	Difficult	Easy	NA
	33. Tasks can be performed in a straightforward manner	Never	Always	NA
	34. Number of steps per task	Too many	Just right	NA
	35. Steps to complete a task follow a logical sequence	Never	Always	NA
	36. Feedback on the completion of the steps	Unclear	Clear	NA

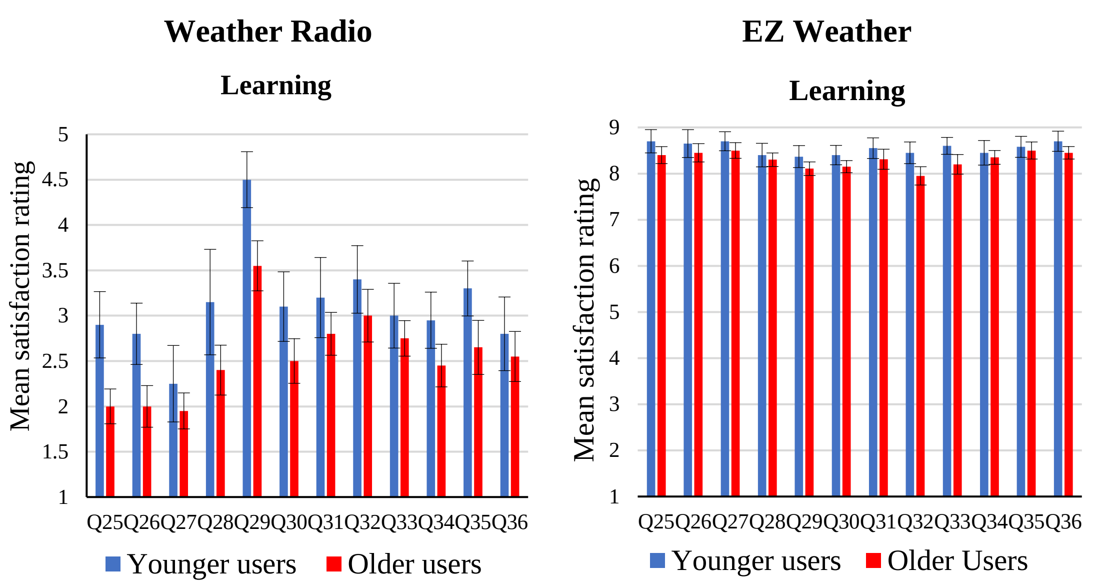


Figure 59. Users' mean satisfaction ratings for the learning category on Weather Radio and EZ Weather

Figure 59 shows that in comparison to Weather Radio, users reported that it was extremely easy to get started and learn using EZ Weather with high efficiency. Moreover, users seemed to be very satisfied with the number and sequence of steps needed to complete each of EZ Weather tasks. One user stated: *"It's very helpful that one single operation completes my task, like the weather forecast task."* They also liked the feedback they received at the completion of each task. Exploration of features and remembering names and commands were also easy for both age groups.

Table 18. QUIS system capabilities & multimedia questions

	Question	1 2 3 4 5 6 7 8 9		
System Capabilities & Multimedia	37. Correcting your mistakes	Difficult	Easy	NA
	38. Correcting typos	Complex	Simple	NA
	39. Ability to undo operations	Inadequate	Adequate	NA
	40. Ease of operation depends on your level of experience	Always	Never	NA
	41. You can accomplish tasks knowing only a few commands	With difficulty	Easily	NA
	42. You can use features/ shortcut	With difficulty	Easily	NA
	43. Colors used are	Unnatural	Natural	NA
	44. Amount of colors available	Inadequate	Adequate	NA

With respect to EZ Weather capabilities, even though many users did not commit any mistakes during the experiment, they reported that it was easy to correct mistakes and typos. Users also liked their ability to use shortcuts for performing or undoing operations. More importantly, they stated that users with any experience level could easily and always accomplish their tasks. The color choices of both weather apps seemed to be adequate and natural for most users, with higher satisfaction for those of EZ Weather. Overall, though both age groups were more satisfied with EZ Weather's

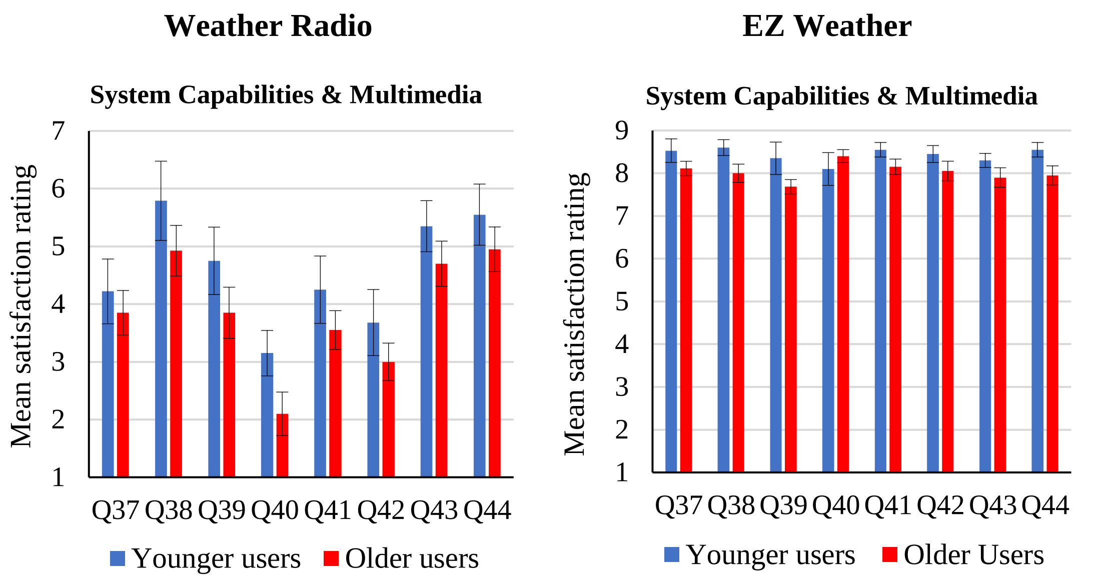


Figure 60. Users' mean satisfaction ratings for the system capabilities & multimedia category on Weather Radio and EZ Weather

system capabilities category than that with Weather Radio, the difference didn't seem to be as significant as it's on the other categories.

Correlations among usability measures

Tables 19 to 22 show the results of the Pearson correlations (*r*) between all metrics of this study. To get a general idea of association, we investigated the average values from all tasks for both age groups on Weather Radio (Tables 19 & 20) and EZ Weather (Tables 21 & 22). The correlations between completion rate and the other metrics for both age groups on EZ Weather were not performed as all users successfully performed the given tasks. The results clearly indicate that there were very strong positive correlations between task time and number of errors and between post-task and post-test satisfaction for both age groups on both apps. It was also shown that each of task time and number of errors was strongly correlated with both post-task and post-test satisfaction ratings in a negative trend. Moreover, the completion rate was negatively correlated with both task time and number of errors, while it was positively and strongly correlated with both post-task and post-test satisfaction ratings.

Table 19. Correlation matrix for younger and on Weather Radio (*r* & (*p*-values))

	Task Time	Number of Errors	Post-Task Satisfaction	Post-Test Satisfaction
Completion Rate	-.687 (.003)	-.647 (.003)	.712 (<.001)	.811 (<.001)
Task Time		.879 (<.001)	-.895 (<.005)	-.898 (<.001)
Number of Errors			-.806 (<.001)	-.844 (<.001)
Post-Task Satisfaction				.937 (<.001)

Table 20. Correlation matrix for older and on Weather Radio (*r* & (*p*-values))

	Task Time	Number of Errors	Post-Task Satisfaction	Post-Test Satisfaction
Completion Rate	-.652 (.002)	-.719 (<.001)	.840 (<.001)	.851 (<.001)
Task Time		.870 (<.001)	-.862 (<.005)	-.877 (<.001)
Number of Errors			-.826 (<.001)	-.913 (<.001)
Post-Task Satisfaction				.922 (<.001)

Table 21. Correlation matrix for younger and on EZ Weather (*r* & (*p*-values))

	Number of Errors	Post-Task Satisfaction	Post-Test Satisfaction
Task Time	.774 (<.001)	-.772 (.005)	-.761 (<.001)
Number of Errors		-.749 (<.001)	-.799 (<.001)
Post-Task Satisfaction			.871 (<.001)

Table 22. Correlation matrix for older and on EZ Weather (*r* & (*p*-values))

	Number of Errors	Post-Task Satisfaction	Post-Test Satisfaction
Task Time	.848 (<.001)	-.843 (.005)	-.884 (<.001)
Number of Errors		-.781 (<.001)	-.823 (<.001)
Post-Task Satisfaction			.900 (<.001)

6.4 Discussion

Overall, the present study found that the UCD weather app, EZ Weather, was noticeably more usable than the representative popular weather app, Weather Radio, in terms of all used evaluation metrics. In addition, both younger and older users appeared to mostly have similar results on all tasks of both weather apps, with slightly higher performance and satisfaction by younger users. These results indicate that apps' interface designs significantly impact end-users' performances and perceptions of apps' usability (either positively or negatively), regardless of age. The results further indicate that employing the UCD approach for apps that include time-critical data, such as weather apps, would result in highly interactive and usable systems.

The results from all metrics imply that prioritizing and structuring critical information as well as using everyday terminology throughout EZ Weather interface substantially helped users to easily interact with the inherent features and perceive them as useful. As an example, all younger and older users were able to accomplish the alert messages task on EZ Weather with considerably higher efficiency and task satisfaction than those on Weather Radio's alert messages feature. Additionally, only a few users made a few errors of cosmetic usability problem category on EZ Weather's alert messages task, while nearly half of the users in each age group made errors due to usability problems of catastrophe category on this task on Weather Radio. The results of the alert messages task further imply that the refined content of pushed alerts would greatly help end-users to correctly perceive and efficiently react to alert threats, especially during imminent weather situations.

The EZ Weather's greater usability is also attributed, in part, to the consideration of efficiency of use heuristic in the design phase. This is supported by the fact that accessing any feature on EZ Weather requires very limited time and number of operational steps. A great example is the weather forecasts feature, which includes all weather forecast information of each saved location within the same screen; with a single step of swiping right or left, users can access weather forecasts of other saved locations. Although a few users could not figure out the swiping functionality, from first trial, all of them successfully completed the weather forecasts task with significantly higher efficiency & task satisfaction, as well as fewer and less severe errors, compared to those on Weather Radio's multiple-screen weather forecasts feature. Another example of efficient features on EZ Weather is the location search feature. Typing familiar locations with effective auto-suggestions enabled both age groups to accomplish the corresponding task in only a few seconds. In contrast, as indicated by the large failure rate and completion time on Weather Radio's location search, navigating the map and pinpointing a location within the limited smart phone screen size was found to be not only an inefficient feature, but also ineffective, especially for older users.

The users' higher performance and satisfaction on EZ Weather over and above Weather Radio might also be related to the minimalist design of EZ Weather. For instance, limiting the alerts controlled by users to the most common and critical ones on EZ Weather enabled users to effectively and efficiently perform the alert settings task with high satisfaction. It is worth mentioning that EZ Weather also allows users to access non-critical alerts (through a representative icon that appears on the affected location's screen) during active alerts. The substantially lower performance and satisfaction on Weather Radio's alert settings task might mainly be due to the large number of alerts and sub-alerts to be navigated and controlled by users (e.g. wind alert alone has 16 sub-alerts).

Another two important heuristics considered in EZ Weather interface design and contributed to its greater usability are affordance and easy recognition. For example, though the map settings task on both apps required interaction with similar steps and function icons, the higher performance and satisfaction on that task of EZ Weather are believed to be attributed to the visibly and intuitively located menus, as well as the appropriately labeled icons. With such factors, the results indicate that executing operations has led to predictable results and followed a logical sequence. In contrast, the invisible elements (e.g. map settings menu bar) and the neither labeled nor standardized icons across smartphone apps may have been among the primary causes for the substantially lower performance and satisfaction on Weather Radio's tasks.

Other usability and design guidelines followed in the design of EZ Weather and may indirectly contributed to the great results included: feedback about the system status (e.g. confirmation messages of executed actions), consistency of the app elements (e.g. settings menu), availability of shortcuts to speed interaction and correction of mistakes, availability of short descriptive information to aid users understanding of the functionality of corresponding features, and use of large text font size as well as high contrast and indicative colors to account for older users' age-related limitations.

The post-test satisfaction findings are in line with the findings from the task-based metrics. The extreme differences between the two apps, in terms of both age groups' satisfaction on all interface criteria, indicate that applying usability guidelines in interface designs not only makes users have high performances, but also makes them satisfied. In addition, the large satisfaction similarity among the two age groups on most of the QUIS survey items implies that the age differences do not significantly impact users' satisfaction levels; what really matters is whether the interface is user-friendly or not.

The strong correlations among all used metrics supplement the findings from previous research such as Joo (2010), suggesting usability metrics are dependent aspects among one another in informing the usability of interfaces. However, a few other studies, such as that by Frøkjær et al. (2000) showed that the dependency of usability metrics relies on whether the tested interface contains highly complex features. If the domain of interest includes complex features, weak correlations are expected; if not, strong correlations are highly possible. In general, we believe all metrics, when integrated, provide great insight towards the usability.

6.5 Chapter Summary

This chapter has presented the last phase (phase 3) in the UCD process: testing the usability of the developed prototype smart phone weather app. To validate the usability of the UCD prototype app, we performed a benchmarking study by comparing the prototype app (EZ Weather) with the popular weather app (Weather Radio). In addition, we considered both younger and older first-time users to examine if the usability of the tested apps would be highly affected by age differences. During the experiment, both age groups performed 5 tasks (location search, weather forecasts, alert messages, map settings, and alert settings) on each of the tested apps. For evaluation, we used task completion rate, number of errors, severity ratings of errors, frequency of errors, task completion time, post-task satisfaction survey, and post-test satisfaction survey.

The results from all measures showed that the enhanced UCD features on EZ Weather: (1) location search (i.e. typing familiar locations with effective auto-suggestions), (2) weather forecasts (i.e. all-inclusive weather forecasts within one screen), (3) alert messages (i.e. use of structured, prioritized, and language-simplified alert messages content), (4) map settings (i.e. use of visible and intuitive map menus), and (5) alert settings (i.e. use of minimalist alert settings), significantly improved both age groups' performance and satisfaction over and above equivalent features on Weather Radio. For correlation tests, it was found that there were very strong correlations among all metrics for both age groups on both apps. Based on these results, employing the UCD approach shows promise in enhancing users' experiences with interfaces of time-critical data and may lead to business success.

Chapter 7. Summary, Recommendation, and Conclusion

7.1 Research Summary

Today, people rely heavily on accessing and receiving weather information, including critical and non-critical information, through smart phone apps (Zabini, 2016). Previous research has not comprehensibly and analytically evaluated the usability of smart phone weather alert apps and whether users of different characteristics (e.g. different age groups) can easily and confidently interact with weather app interfaces. In the present work, we have performed a comprehensive systematic approach for evaluating and improving the usability of smart phone weather apps. In Chapter 3, we performed a mixed methods analysis with the goal of evaluating first-time users' performances on and perceptions about the usability of weather apps and mainly discovering usability problems. In the following three chapters, we employed the user-centered design approach (UCD) with the goal of developing a usable weather app from end-users' perspectives. Each of the three UCD chapters discussed one of the UCD phases depicted in Figure 2. Specifically, based on the observed usability problems and the knowledge attained from the mixed methods analysis, in Chapter 4, we created a set of critical questions and presented them to a sample of weather app users in a focus group setting, to discover more usability problems and primarily learn about users' needs for future usable weather apps. Additionally, Chapter 4 presented general usability heuristics, smart phone app design guidelines, and age-related smart phone best practices. In the following chapter, we developed a prototype smart phone weather app considering users' feedback along with key usability and smart phone app design heuristics, including special heuristics for older users (see Chapter 4). Finally, in Chapter 6, we tested the usability of the developed UCD prototype app to see if it has practically addressed users' needs and suited their different characteristics and challenges.

Summary of the mixed methods findings

The mixed methods analysis revealed that though all users successfully completed the Weather Radio app's search tasks (i.e. location search, alert settings, and map settings), first-time users performed substantially worse than experienced users on all used measures on the alert and map settings tasks. No significant difference was observed between the two groups on both location search approaches: a pin on the map and typing in the app's text bar, with better performance on the typing in the app's text bar approach. Similar to the location search task, no significant difference was observed between the two user groups on both alert messages approaches: original NWS messages and proposed messages, where both first-time and experienced users rated the two approaches similarly, with substantially higher satisfaction ratings on the proposed ones.

The combination of both the traditional and eye tracking metrics provided great quantitative and qualitative insights about the usability problems and the users' cognitive processes and decision-making strategies. Specifically, the used metrics enabled us to discover several usability problems and violations to key usability principles such as visibility, affordance, efficiency of use, and appropriate use of language. For example, lack of visibility and affordance were the main problems for first-time users' poor performance on the map settings task. The users struggled a lot finding the function required to access the map settings menu; this was included in an extended navigation menu. Due to the limited smart phone screen size, this menu was invisibly located and required two prior counterintuitive steps (see Figure 7 for visual explanation). This poor performance was explained by the large completion time on this task and later justified by the eye tracking measures. The eye tracking measures showed that most of the eye fixation counts and durations as well as the unnecessary eye transitions occurred before performing the first correct step required for this task: finding the map settings menu. The lack of affordance was also one of the factors for first-time users' poor performance, as all the app's icons (e.g. map settings icons) were unlabeled and users could not intuitively figure out their functionality. The eye tracking measures supported this claim, as the users frequently scanned and randomly tapped the menu icons.

Summary of the first UCD phase findings

This section contains the findings from the first UCD phase: focus group interviews, general usability heuristics, smart phone app design guidelines, and smart phone best practices regarding age-related limitations. The focus groups findings revealed that the participants were mostly looking for weather app interfaces that ease their cognitive processes and lessen their cognitive loads. They repeatedly expressed a need for intuitive interfaces, minimum, easy to understand, and nicely visualized information. The participants also voiced a need for effective interfaces by prioritizing critical (e.g. alert messages) and frequently used (e.g. weather forecasts) features on the interface. Moreover, they believed that a reduced number of operational steps would substantially enhance the efficiency of weather app interfaces, leading to a quick information accessibility, especially during imminent severe weather conditions.

The user needs elicited in the focus groups were mostly in line with the general usability heuristics proposed by Nielsen (1995a) and discussed thoroughly in Chapter 4. Examples of the general heuristics are speaking the user's language, using consistent, flexible, easy to recognize, and efficient interfaces, and using minimalist design. Even though most usability heuristics are needed for users of different age groups, special heuristics related to older users' limitations were also highlighted in Chapter 4 and believed to not negatively impact younger users' interactions, such as high color contrasts, relatively large font size and icons, and large spaces between items. Finally, we presented key smart phone app heuristics such as labeling menus and icons, associating menu options with helpful descriptive information, using effective location search index (e.g. auto suggestions), and using fixed navigation menus.

Summary of the second UCD phase

The second UCD phase was discussed in Chapter 5. In Chapter 5, we showed the structure and content of the developed smart phone prototype weather app (EZ Weather). By using visual and text explanation, we showed examples of the addressed heuristics in EZ Weather as well as the steps needed to access and control all features. In EZ Weather, we addressed all user needs and carefully considered usability heuristics and smart phone app interface guidelines such as those illustrated in Chapter 4. The

UCD prototype app also paid specific attention to the age-related limitations and considered the heuristic guidelines for older users.

Summary of the third UCD phase

Chapter 6 covered the third and last UCD phase, which was testing the usability of EZ Weather. Overall, the analysis showed that both younger and older users performed substantially better on EZ Weather compared to their performance on Weather Radio. Not only that, but they also perceived it as noticeably more usable than Weather Radio. The analysis also showed that no significant difference was observed between the two age groups on all tasks of both apps, in terms of performance and satisfaction levels, with slightly better performance and satisfaction by younger users. The enhanced UCD features in EZ Weather that contributed to these results are: (1) location search (i.e. typing familiar locations with effective auto-suggestions), (2) weather forecasts (i.e. all-inclusive weather forecasts within one screen), (3) alert messages (i.e. use of structured, prioritized, and language-simplified alert messages content), (4) map settings (i.e. use of visible and intuitive map menus), and (5) alert settings (i.e. use of minimalist alert settings). The results indicate that apps' interface designs substantially impact end-users' performances and perceptions of apps' usability (either positively or negatively), regardless of age. The results further indicate that employing the UCD approach for apps that include time-critical and/or life-saving data, such as weather apps, would result in highly interactive and usable systems.

7.2 Research Contribution

The overall contribution of this dissertation has been to generate critical information and practical evidences for smart phone weather alert app developers in order to build greatly usable interfaces. As illustrated in the literature review (see Chapter 2), the smart phone app availability and usage are astonishingly increasing (Stacy, 2017). For example, weather information is utilized on smart phone apps more than on all other information sources (Zabini, 2016). The smart phone weather apps are among the seven most used smart phone app categories in the United States (Statista, 2018a) with more than 8000 weather apps available in the iTunes app store alone, as of August 2018 (iTunes, 2018). Hence, the usability investigation of this area is extremely critical, as users need to easily and efficiently access the required information, especially during severe weather conditions such as tornados, hurricanes, and floods, in order to perform life-saving actions.

Previous research has not well established knowledge regarding the features of smart phone weather apps. Among the few available studies on weather apps are studies by Singhal (2011) and Alluri (2012). However, both studies tested the native (originally built-in) weather apps that include only basic weather forecasts, where the findings may not be sufficient and/or compliant with the actual demand in the field. Even though Drogalis et al. (2015) later evaluated the usability of one of the most advanced and popular weather apps, they did not consider a benchmark approach that links the users' performance to a standard data in order to logically determine whether the user's performance was satisfactory or poor. In addition, the sample size (6 participants) used in their study was insufficient to generalize the results.

To promote users' interactions with weather alert apps and account for the limitations in previous research, this dissertation first presents a comprehensive analytical evaluation of the usability of smart phone weather apps using traditional metrics (e.g. completion time and surveys) and advanced eye tracking-based measures. With this evaluation, we clearly identify the usability problems that might hinder performing life-saving actions.

Second, the dissertation uses a UCD approach with the goal of enhancing users' experiences with weather apps. The UCD approach employed in this dissertation starts with qualitatively (using focus group interviews) identifying end-user exact needs for future weather apps. Following that, a prototype weather app is developed by carefully considering user needs voiced in the focus groups as well as following the usability and smart phone app design heuristics. When developing the prototype, age-related limitations and recommended heuristics were also considered in order to accommodate the needs of this important age group as well as produce a user-friendly interface for users of all ages. In the final UCD phase, a usability evaluation of the developed prototype app is performed in order to validate the usability of the app and what extent usability problems and user needs are practically addressed.

The findings from the work in this dissertation are believed to substantially fill several gaps in the existing literature regarding the usability of smart phone weather alert apps. The acquired knowledge is also claimed to help smart phone app developers design user-friendly interfaces, especially those concerned with time-critical and life-saving apps. Such enhanced apps are not only to the users' benefit, but also to businesses as they are expected to experience growth and gain more credibility.

7.3 Practical Recommendation

From the findings of all conducted studies, five feature-wise recommendations were generated for future designs of weather app interfaces. First, weather app developers should practically employ the UCD approach and think of formatting the alert messages from the end-user perspective. Specifically, the alert messages need to go through a filtration process before they reach the end-user. Examples of the actions that may need to be implemented in the filtration process are: 1) removing the geographical area codes for the locations under alerts and including only location names, 2) including information only about the user's saved location and the nearby under-alert locations on the main message screen; information regarding all other affected locations could be included on a secondary screen that can be accessed once the user taps a certain option/icon from the main message screen, 3) hierarchically structuring the alert content based on importance, and 4) using an appropriate language that considers the differences in users' characteristics.

Second, developers should consider enhancing the location search capabilities by allowing users to add an exact location name/address; a user gets weather information with respect to their exact location. For example, the Weather Radio app sends alert notifications to end-users only if their exact saved location falls within the watch/warning box specified by the NWS. On the other hand, as most of the currently available weather apps only use general location search methods (e.g. zip, county, and/or city), they may not provide weather forecasts and send alerts that reflect the actual weather condition at the user's exact location. In other words, if only a tiny part of the zip, city, or county is forecasted to be affected by a specific weather condition, the app would generalize it to the whole area, leading to false alarms. As discussed in the focus groups (see Chapter 4), though most of the participants had not used an exact location search feature, they stated that using it in the future would make them feel comfortable about the accuracy of the weather forecasts. In the quote of one participant:

> "I think adding my exact location will let me get accurate weather information, because sometimes I get wrong forecasts as my app forces me to add either the city name or the zip code. The app says it is raining at my school, but in fact it is not. It may be raining in another part even though both locations are within the same zip code area."

The exact location search feature on weather apps should also be more efficient. As was proven in this dissertation, typing the exact location in the search field was substantially more efficient and satisfactory for all users than searching the map.

Third, as the participants stated in the focus group interviews that the weather forecasts and alert messages are the most important and/or frequently accessed features, developers should give more priority to these two features by facilitating users' access to and interaction with them. As an example, the UCD prototype app, EZ Weather, included each saved location, along with its daily and hourly weather forecasts, within a single home screen (see Figure 48). This way allows users to access the weather forecasts as soon as they open the app rather than going through multiple navigational steps. The alert message feature can be accessed from the same screen of each saved location through a representative standardized weather alert icon placed in an appropriate position (see Figure 49). This icon appears only when there is an active alert for the saved location.

Fourth, developers should design visible and intuitive elements, in a way that these elements and actions always lead to predictable results, throughout the entire interface. This is recommended here as the participants struggled a lot finding elements on the map settings feature of Weather Radio, due to the violations to these principles (see Chapters 3 and 6). In addition, the focus group participants repeatedly emphasized the need for having weather app interfaces with minimal cognitive load, with several shared negative examples on the radar map feature in weather apps. To enhance interfaces' visibility and affordance on the radar map feature, developers should clearly indicate the process of accessing each saved location's map, make the map menu in a visible location, and use representative icons with appropriate and intuitive labels.

Last, the alert settings should be optimized. Weather app developers should follow the minimalist design heuristic by limiting the number of alerts to the most critical and common ones (e.g. those in EZ Weather) for user control, instead of displaying an overwhelming number of alerts and sub-alerts (e.g. those in Weather Radio). However, to give users the flexibility to access non-critical weather alerts (e.g. wind watch), developers should enable users to access the related messages without a need for controlling their settings. For example, EZ Weather displays the messages of such alerts once a representative alert icon, with an attached number indicating the number of active alert messages, is tapped; this icon is placed on the affected location's screen (see Figure 49).

Other indirect feature-wise recommendations based on the best practices for smart phone app designs, usability heuristics, as well as the findings of the work in this dissertation should:

- Add concise descriptive information to further explain menu options.
- Use swiping screens, especially for apps that include structured and focused content.
- Provide clear and concise confirmation messages to assure the user about the completed action.
- Keep the navigation menu bars fixed throughout the entire app in order to speed the accessibility of functions as well as recoverability from unwanted actions or slips.
- Keep all related elements, menus, and layouts consistent so that users can intuitively understand the presented information and react accordingly.

There are also other important age-related recommendations as follows:

- Reduce the number of elements (e.g. icons) on each screen and increase their sizes and spacing between them.
- Make text font size no less than 12 point.
- Use bright screens and high color contrasts (e.g. black text on white screens).
- Use the minimum amount of necessary information on each screen.

7.4 Limitations and Future Research

One of the limitations of the work in this dissertation is that only the iOS, along with its users were considered for all conducted studies. As other smart phone operating systems (e.g. Android) mostly have unique structures and design specifications and guidelines (Li et al., 2012), the usability findings of our work might not be completely or sufficiently generalizable to users of such systems. For a future research, different operating systems should be considered for smart phone weather app usability evaluation in order to accommodate the needs of different user populations.

Another potential limitation, all our conducted studies arbitrarily ran experiments on an iPhone 6, which has a medium screen size; there are a variety of screen sizes. A future study could test if the smartphone screen size would produce different results. In addition, a future study regarding weather alert systems could test the effect of different screen sizes on different platforms such as iPads, tablets, and desktop computer displays. The findings from this research may further determine the importance of the screen size factor for usability evaluation.

A third limitation is that although the eye tracking study (see Chapter 3) identified multiple usability problems associated with the interface designs of smart phone weather alert apps with particular focus on the Weather Radio app and discussed several usability principles, the testing environment may have not fully simulated the real-world situation. Specifically, the placement of the cell phone on the monitor was not simulating the users' typical use in everyday life. However, as a remote eye tracker with cameras attached under a monitor was only available for this study, it was not possible to have the users hold the cell phone by hand while interacting with the Weather Radio app; the eye tracker would have been blocked by users' hands and the cell phone and affected by the users' head down bending. In future research, a wireless head-mounted eye tracker or eyeglasses tracker could be used so that it captures users' eye movements accurately as well as provides a realistic environment through having the cell phone in hand. In particular, the head-mounted eye tracker is recommended to be used in future research as it was previously employed to gauge users' interaction with smart phone apps and accurately captured their eye movements (Chynał, 2012). In addition, the eye tracking findings of this study show the possibility for developing a future predictive method in a controlled and sophisticated manner, to know whether the user is a first-time or experienced from their eye movements.

Last, the usability problems were objectively identified from users' interaction with only one popular smart phone weather app. Even though the qualitative focus group interviews (see Chapter 4) uncovered several other usability problems, based on usage of eleven different popular weather apps, a future study could consider objective evaluation of multiple smart phone weather apps to account for the limitation of the single app usage as well as support the subjective evaluation.

7.5 Conclusion

From a theoretical perspective, this work adds breadth to the literature regarding the usability and UCD investigations of smart phone weather apps. In addition, our current work clearly indicates the importance of considering the observed usability problems and end-users' needs in interface designs, particularly those containing time-critical data, and how they greatly contribute to enhancing user performance and satisfaction. Furthermore, this research contributes to knowledge about the ways in which usability heuristics and smart phone app design guidelines need to be followed and prioritized to meet the needs of smart phone weather apps' end-users. Moreover, the research illustrates the importance of addressing age-related limitations and considering recommended heuristics in smart phone app designs to accommodate the needs of older users. Consideration of all these heuristics and guidelines are believed to lead to both customer trust and business success.

Even though this work thoroughly and specifically explored the usability and UCD in weather apps, the findings and recommendations may provide insight in other smart phone app categories that involve time-critical and/or life-saving information, such as those of emergency medicine.

References

Abras, C., Maloney-Krichmar, D., & Preece, J. (2004). User-centered design. Bainbridge, W. Encyclopedia of Human-Computer Interaction. Thousand Oaks: Sage Publications, 37(4), 445-456.

Abras, C., Maloney-Krichmar, D., Preece, J. (2004) User-Centered Design. In *Bainbridge, W. Encyclopedia of Human-Computer Interaction*. Thousand Oaks, CA: Sage Publications.

Adcock, A. (2000). Effects of cognitive load on processing and performance. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download? doi=10.1.1.90.387&rep=rep1&type=pdf.

Ahmadi, H., & Kong, J. (2008). Efficient web browsing on small screens. In *Proceedings of the working conference on Advanced visual interfaces* (pp. 23-30). Napoli, Italy: ACM.

Ahmadi, H., & Kong, J. (2008, May). Efficient web browsing on small screens. In *Proceedings of the working conference on Advanced visual interfaces* (pp. 23-30). New York, USA: Association for Computing Machinery.

Albert, W., & Tullis, T. (2013). *Measuring the user experience: collecting, analyzing, and presenting usability metrics*. Amsterdam, Netherland: Elsevier.

Al-khomsan, M., Al-Arjan, A., Al-Amro, A., & Al-Nafjan, K. (2015). Usability evaluation of Twitter on Android platform for elderly Arab users using morae evaluation tool. In *Proceedings of the Internet Technology and Secured Transactions 10th International Conference for Internet Technology and Secured Transactions* (pp. 446-447). London, UK: Institute of Electrical and Electronics Engineers.

Alluri, A. (2012). *Usability testing of android applications* (Master's thesis). Retrieved from http://sdsu-dspace.calstate.edu/bitstream/ handle/10211.10/3556/ Alluri Aruna.pdf? sequence=1.

Alshehri, F. & Freeman, M. (2012). Methods for usability evaluations of mobile devices. In *Proceedings of the 23rd Australian Conference on Information Systems* (pp. 1-10). Geelong, Australia: ACIS.

Anthony, T. (2012). Finger-Friendly Design: Ideal Mobile Touchscreen Target Sizes. Retrieved from https://www.smashingmagazine.com/2012/02/finger-friendly-design-ideal-mobile-touchscreen-target-sizes/.

Argyle, E. (2015). *Supporting situation awareness and decision making in weather* *forecasting* (Doctoral dissertation). Retrieved from https://shareok.org/handle/11244/34617.

Argyle, E., Gourley, J., Clark, R., Flamig, Z., Gutierrez, M., Erlingis, J., Martinaitis, S., Smith, B., Ling, C. (2015). Communicating threat level for flash floods in the hazardous weather testbed hydrology experiment 2014. In *Proceedings of the 3rd Conference on Weather Warnings and Communication.* Raleigh, NC: American Meteorological Society.

Argyle, E., Gourley, J., Flamig, Z., Hansen, T., & Manross, K. (2017). Toward a user-centered design of a weather forecasting decision-support tool. *Bulletin of the American Meteorological Society*, *98*(2), 373-382.

Baharuddin, R., Singh, D., & Razali, R. (2013). Usability dimensions for mobile applications: a review. *Research Journal of Applied Sciences, Engineering and Technology, 5*(6), 2225-2231.

Ball, L., & Bothma, T. (2017). The Importance of usability evaluation when Developing Digital tools for a library-a Case Study. In *Proceedings of BOBCATSSS* (pp. 181 – 187). Tampere, Finland: BOBCATSSS.

Barendregt, M., Bekker, M. (2003). Empirical evaluation of usability and fun in computer games for children. In *Proceedings of Human-Computer Interaction* (pp. 705-708): IOS Press

Barros, A., Leitão, R., & Ribeiro, J. (2014). Design and evaluation of a mobile user interface for older adults: Navigation, interaction and visual design recommendations. *Procedia Computer Science, 27*(1), 369-378.

Bitterman, N., & Shalev, I. (2004). The silver surfer: Making the internet usable for seniors. *Ergonomics in Design: The Quarterly of Human Factors Applications, 12*(1), 24-28.

Bojko, A., & Schumacher, R. (2008). Eye tracking and usability testing in form layout evaluation. In *Proceedings of the 38th International Symposium of Business Forms Management Association* (pp. 1–13). Las Vegas, USA: BFMA.

Bourie, P., Dresch, J., & Chapman,H. (1997). Usability evaluation of an on-line nursing assessment. In *Proceedings of the AMIA Annual Fall Symposium* (pp. 406-415). Nashville, TN: American Medical Informatics Association.

Bowman, D., Gabbard, J., & Hix, D. (2002). A survey of usability evaluation in virtual environments: classification and comparison of methods. *Presence: Teleoperators & Virtual Environments, 11*(4), 404-424.

Boyatzis, R. (1998). *Transforming qualitative information: Thematic analysis and code development*. Thousand Oaks, CA: Sage Publications.

Braun, V. & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77-101.

Braun, V., Clarke, V., & Terry, G. (2014). Thematic Analysis. In P. Rohleder & A. C. Lyons (Eds.), *Qualitative Research in Clinical and Health Psychology*. China: Palgrave Macmillan.

Brown, W., Yen, P., Rojas, M., & Schnall, R. (2013). Assessment of the Health IT Usability evaluation model (Health-ITUEM) for evaluating mobile health (mHealth) technology. *Journal of biomedical informatics, 46*(6), 1080-1087.

Cáliz, D., & Alamán, X. (2014). Usability evaluation method for mobile applications for the elderly: a methodological proposal. In *Proceedings of the International Workshop on Ambient Assisted Living 6th Annual Meeting* (pp. 252-260). Belfast, UK: Springer.

Caplan, S. (1990). Using focus group methodology for ergonomic design. *Ergonomics, 33*(5), 527-533.

Charness, N., Kelley, C., Bosman, E., & Mottram, M. (2001). Word-processing training and retraining: Effects of adult age, experience, and interface. *Psychology and Aging, 16*(2), 110–127.

Cheng, S. (2011). The research framework of eye-tracking based mobile device usability evaluation. *Proceedings of the 1st international workshop on pervasive eye tracking & mobile eye-based interaction* (pp. 21-26). New York, USA: Association for Computing Machinery.

Chervinskaya, K., & Wasserman, E. (2000). Some methodological aspects of tacit knowledge elicitation. *Journal of Experimental and Theoretical Artificial Intelligence, 12*(1), 43–55.

Chin, P., Diehl, A. and Norman, L. (1988). Development of an instrument measuring user satisfaction of the human-computer interface. In *Proceedings of the Special Interest Group on Computer-Human Interaction conference on human factors in computing systems 6th Annual Meeting*, (pp. 213-218), New York, NY: Association for Computing Machinery.

Chisnell, D. E., Redish, J. C. G., & Lee, A. M. Y. (2006). New heuristics for understanding older adults as web users. *Technical Communication, 53*(1), 39-59.

Chittaro, L., Zuliani, F., & Carchietti, E. (2007). Mobile devices in emergency medical services: user evaluation of a PDA-based interface for ambulance run reporting. In Proceedings of *International Workshop on Mobile Information Technology for Emergency Response* (pp. 19-28). Berlin, Heidelberg: Springer.

Chynał, P., Szymański, J., & Sobecki, J. (2012). Using eye tracking in a mobile apps usability testing. *Intelligent Information and Database Systems*, *7193*(1), 178-186.

Clamann, M., & Kaber, D. (2004). Applicability of usability evaluation techniques to aviation systems. *The international journal of aviation psychology*, *14*(4), 395-420.

Constantine, L., & Lockwood, L. (1999). *Software for use: a practical guide to the models and methods of usage-centered design*. Boston, MA: Pearson Education.

Cooke, L. (2006). Is eye tracking the next step in usability testing? In *Proceedings of IEEE International Professional Communication Conference* (pp. 236-242). Washington DC, USA: IEEE.

Czaja, S., Charness, N., Fisk, A., Hertzog, C., Nair, S., Rogers, W., & Sharit, J. (2006). Factors predicting the use of technology: Findings from the Center for Research and Education on Aging and Technology Enhancement (CREATE). *Psychology and Aging, 21*(2), 333–352.

Dix, A., Finlay, J., & Abowd, G. (2004). *Human-computer interaction (3rd ed.)*. Harlow, UK: Pearson Education.

Doglas, S. (2017). Seven usability heuristics that All UI designers should know. Retrieved from https://usabilitygeek.com/usability-heuristics-ui-designers-know/.

Donker, A., & Reitsma, P. (2007). Young children's ability to use a computer mouse. *Computers & Education*, *48*(4), 602-617.

Drogalis, T., Keyes, E., & Dhyani, L. (2015). The Weather Channel application usability test. Retrieved from http://elkeyes.com/wp-content/uploads/2015/05/WeatherChannelUTResults.pdf.

Ehmke, C., & Wilson, S. (2007). Identifying web usability problems from eye-tracking data. In *Proceedings of the 21st British HCI Group Annual Conference on People and Computers: HCI* (pp. 119-128). Swindon, UK: British Computer Society.

Elzakker, C., Delikostidis, I., & Oosterom, P. (2008). Field-based usability evaluation methodology for mobile geo-applications. *The Cartographic Journal*, *45*(2), 139-149.

EMarketer (2015). App Marketing: Fighting for Downloads and Attention in a Crowded Market. Retrieved from http://www.nanigans.com/blog/cross/usu/3-key-takeaways-from-emarketers-2015-app-marketing-report/.

Encarnaçao, J., Foley, J., Bryson, S., Feiner, S., & Gershon, N. (1994). Research issues in perception and user interfaces. *IEEE Computer Graphics and Applications, 14(*2), 67-69.

Faulkner, L., & Wick, D. (2005). Cross-user analysis: Benefits of skill level comparison in usability testing. *Interacting with Computers*, *17*(6), 773-786.

Feng, F., Liu, Y., & Chen, Y. (2017). A computer-aided usability testing tool for in-vehicle infotainment systems. *Computers & Industrial Engineering, 109*(3), 313-324.

Fisk, A., Rogers, W., Charness, N., Czaja, S., & Sharit, J. (2009). *Designing for older adults: Principles and creative human factors approaches*. Boca Raton, FL: CRC Press.

Foraker, A. (2018). Introduction to User-Centered Design. Usability first. Retrieved from http://www.usabilityfirst.com/about-usability/introduction-to-user-centered-design/.

Fu, J. (2016). Usability evaluation of software store based on eye-tracking technology. In *Proceedings of the Information Technology, Networking, Electronic and Automation Control Conference* (pp. 1116-1119). Chongqing, China: IEEE.

García-Peñalvo, F., Conde, M., & Matellán-Olivera, V. (2014). Mobile apps for older users–The development of a mobile apps repository for older people. In *Proceedings of the International Conference on Learning and Collaboration Technologies 1st Annual Meeting* (pp. 117-126). Heraklion, Greece: Springer.

Garrett, J. (2010). *Elements of user experience, the: user-centered design for the web and beyond*. London, UK: Pearson Education.

Georgsson, M., & Staggers, N. (2015). Quantifying usability: an evaluation of a diabetes mHealth system on effectiveness, efficiency, and satisfaction metrics with associated user characteristics. *Journal of the American Medical Informatics Association*, *23*(1), 5-11.

Gerardo, J. (2007). *The effectiveness of novice users in usability testing* (Master's thesis). Retrieved from https://www.duo.uio.no/ bitstream/handle/ 10852/

9681/Gerardo.pdf?sequence=1&isAllowed=y.

Geven, A., Sefelin, R., & Tscheligi, M. (2006). Depth and breadth away from the desktop: the optimal information hierarchy for mobile use. In *Proceedings of the 8th conference on Human-computer interaction with mobile devices and services* (pp. 157-164). Helsinki, Finland: Association for Computing Machinery.

Gladkiy, S. (2018). User-Centered Design: Process and Benefits. Retrieved from https://uxplanet.org/user-centered-design-process-and-benefits-fd9e431eb5a9.

Goldberg, J., & Kotval, X. (1999). Computer interface evaluation using eye movements: methods and constructs. *International Journal of Industrial Ergonomics*, *24*(6), 631-645.

Goldberg, J., & Wichansky, A. (2003). Eye tracking in usability evaluation: A practitioner's guide. In *the Mind's Eye* (pp. 493-516). Amsterdam, Netherlands: Elsevier.

Gómez, R., Caballero, D., & Sevillano, L. (2014). Heuristic evaluation on mobile interfaces: a new checklist. *The Scientific World Journal*, *19*(3), 39-58.

Goodman, J., Dickinson, A., & Syme, A. (2004). Gathering requirements for mobile devices using focus groups with older people. In *Proceedings of Designing a More Inclusive World* (pp. 81-90). London, UK: Springer.

Gove, J. (2016). Principles of Mobile App Design: Engage Users and Drive Conversions. Retrieved from https://www.thinkwithgoogle.com/marketing-resources/experience-design/principles-of-mobile-app-design-engage-users-and-drive-conversions/.

Griffiths, S. (2015). Mobile App UX Principles. Google. Retrieved from https://storage.googleapis.com/think-emea/docs/article/ Mobile App UX Principles.pdf.

Han, S. H., Kim, K. J., Yun, M. H., Hong, S. W., & Kim, J. (2004). Identifying mobile phone design features critical to user satisfaction. *Human Factors and Ergonomics in Manufacturing*, *14*(1), 15–29.

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

Hawthorn, D. (2000). Possible implications of aging for interface designers. *Interacting with Computers*, *12*(5), 507-528.

Holland, C., Garza, A., Kurtova, E., Cruz, J., & Komogortsev, O. (2013). Usability evaluation of eye tracking on an unmodified common tablet. In *Proceedings of the CHI'13 Extended Abstracts on Human Factors in Computing Systems* (pp. 295-300). New York, USA: Association for Computing Machinery.

Holzinger, A., & Errath, M. (2007). Mobile computer Web-application design in medicine: some research based guidelines. *Universal Access in the Information Society*, *6*(1), 31-41.

Hussain, A., & Kutar, M. (2009). Usability metric framework for mobile phone application. In *Proceedings of the 10th Annual Post-Graduate Symposium on The Convergence of Telecommunications, Networking and Broadcasting* (pp. 34-39). London, UK: Bournemouth University.

Hussain, A., & Mkpojiogu, E. (2015). An application of ISO/IEC 25010 standard in the quality-in-use assessment of an online health awareness system. *Jurnal Teknologi (Sciences & Engineering)*, *77*(5), 9-13.

Hussain, A., Mkpojiogu, E., & Hussain, Z. (2015). Usability evaluation of a web-based health awareness portal on Smartphone devices using ISO 9241-11 model. *Jurnal Teknologi*, *77*(4), 1-5.

Hussain, A., Mkpojiogu, E., Musa, J., & Mortada, S. (2017). A user experience evaluation of Amazon Kindle mobile application. In *Proceedings of the 2nd International Conference on Applied Science and Technology* (pp. 11-18). Kedah, Malaysia: AIP Publishing.

Hussain, A., Razak, H., & Mkpojiogu, E. O. (2017). The perceived usability of automated testing tools for mobile applications. *Journal of Engineering, Science and Technology (JESTEC)*, *12*(4), 89–97.

Ilyas, M. (2012). A study of web accessibility barriers for older adults, and heuristics evaluation of email websites based on web accessibility heuristics for older adults by AARP. *Journal of Emerging Trends in Computing and Information Sciences, 3*(5), 806-813.

Inostroza, R., Rusu, C., Roncagliolo, S., & Rusu, V. (2013). Usability heuristics for touchscreen-based mobile devices: update. In *Proceedings of the 2013 Chilean Conference on Human-Computer Interaction* (pp. 24-29). Temuco, Chile: Association for Computing Machinery.

Institute of Electrical and Electronics Engineers Computer Society. (1993). IEEE Standard for a Software Quality Metrics Methodology (IEEE 1061). Retrieved from https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=237006.

International Organization for Standardization. (2001). *Software engineering - Product quality - Part 1: Quality model (IEC 9126-1)*. Geneva, Switzerland: International Organization for Standardization.

International Organization for Standardization. (2010). *Human-centered design for interactive systems (ISO 9241-210:2010)*. Retrieved from https:// www .iso. org /obp /ui/ #iso: std:iso: 9241: -210:en.

International Standards Organization. (1997). *Ergonomic requirements for office work with visual display terminals (VDTs)-Part 11 (ISO 9241): Guidance on usability*. Geneva, Switzerland: International Standards Organization.

iTunes. (2018). Popular weather apps. Retrieved from https://itunes.apple.com/us/genre/ios-weather/id6001?mt=8.

Jacks, E., Pirring, A., Soroka D., Hawkins, M., Allen, A., Mullusky, M, Klockow, K., Esco, G., Girardi, L. (2018). NWS Hazard Simplification Project: A Roadmap for Change. In *Proceedings of* the 6th *Symposium on Building a Weather-Ready Nation: Enhancing Our Nation.* Austin, TX: American Meteorological Society.

Jacob, R., & Karn, K. (2003). Eye tracking in human-computer interaction and usability research: Ready to deliver the promises. In *The mind's eye* (pp. 573-605).

John, B. E., & Kieras, D. E. (1996). Using GOMS for user interface design and evaluation: Which technique? *ACM Transactions on Computer-Human Interaction (TOCHI)*, *3*(4), 287-319.

Jones, W., Drake, C., Mack, D., Reeder, B., Trautner, B., & Wald, H. (2017). Developing Mobile Clinical Decision Support for Nursing Home Staff Assessment of Urinary Tract Infection using Goal-Directed Design. *Applied Clinical Informatics, 8*(2), 632-650.

Joo, S. (2010). How are usability elements-efficiency, effectiveness, and satisfaction-correlated with each other in the context of digital libraries? In *Proceedings of the American Society for Information Science and Technology 73rd Annual Meeting* (pp. 1-2). Hoboken, NJ: John Wiley & Sons.

Kaplan, A. M. (2012). If you love something, let it go mobile: Mobile marketing and mobile social media 4x4. *Business Horizons, 55*(2), 129-139.

Kaufman, J. (2016). Principles of mobile app designs. Retrieved from https://www.apteligent.com/wp-content/uploads/2016/07/00-Whitepaper-Principles-of-App-Design-V1-1.pdf.

Khamaj, A., Shehab, R., and Kang, Z. (2017). Efficiency and distraction evaluations of motorcycles mounted infotainment systems and controls. In *Proceedings of the 6th Annual World Conference of the Society for Industrial and Systems Engineering* (pp. 169-175). Herndon, VA: SISE.

Khamaj, A., and Kang, Z. (2018). Usability evaluation of mobile weather hazard alert applications. *Industrial and Systems Engineering Review, 6(1), 21-40.*

Khanum, M., & Trivedi, M. (2012). Take care: a study on usability evaluation methods for children. *International Journal of Advanced Research in Computer Science*, *3*(2), 101-105.

Khawaji, M. (2017). *Overcoming challenges in smart phone use among older adults in Saudi Arabia* (Doctoral dissertation). Retrieved from https://lib.dr.iastate.edu /cgi/ viewcontent.cgi? article = 7162 & context= etd.

Kjeldskov, J., Skov, M., & Stage, J. (2005). Does time heal? A longitudinal study of usability. In *Proceedings of the 17th Australia conference on Computer-Human Interaction: Citizens Online: Considerations for Today and the Future* (pp. 1-10). Canberra, Australia: Computer-Human Interaction Special Interest Group (CHISIG) of Australia.

Krueger, R. & Casey, M. (2002). Designing and conducting focus group interviews. *Social Analysis, Selected Tools and Techniques*, *21*(1), 4-23.

Kurniawan, S. (2008). Older people and mobile phones: A multi-method investigation. *International Journal of Human-Computer Studies, 66*(12), 889-901.

Kurniawan, S., & Zaphiris, P. (2005). derived web design guidelines for older people. In *Proceedings of the 7th international ACM SIGACCESS conference on Computers and accessibility* (pp. 129-135). New York: ACM.

Lack, R. (2007). The importance of user-centered design: Exploring findings and methods. *Journal of Archival Organization*, *4*(1-2), 69-86.

Lawton, P. (1990). Aging and performance of home tasks. *Human factors*, *32*(5), 527-536.

Lee, J., Niko, D., Hwang, H., Park, M., & Kim, C. (2011). A GIS-based design for a smartphone disaster information service application. In *Proceedings of* the *Computers, Networks, Systems and Industrial Engineering International Conference* (pp. 338-341). Jeju Island, South Korea: IEEE.

Lettner, F., & Holzmann, C. (2011). Usability evaluation framework. In Proceedings of the *International Conference on Computer Aided Systems Theory* (pp. 560-567). Berlin, Germany: Springer.

Leung, R., Findlater, L., Mcgrenere, J., Graf, P., & Yang, J. (2010). Multi-layered interfaces to improve older adults' initial learnability of mobile applications. *Transactions on Accessible Computing, 3*(1), 1-30.

Li, F., Wang, Y., Wu, J., Jiang, K., & Liu, W. (2012). Mobile OS architecture trends. *Intel Technology Journal*, *16*(4), 178-198.

Liu, S. (2012). *Technology Acceptance Model for Determining the Effects of Age, Usability, and Content on Mobile Application Usage* (Master's thesis). Retrieved from https://etd.ohiolink.edu/!etd.sendfile? accession= ohiou 1339700970&disposition=inline.

Lizano, F., Sandoval, M., Bruun, A., & Stage, J. (2013). Is usability evaluation important: The perspective of novice software developers. In *Proceedings of the 27th International BCS Human Computer Interaction Conference* (pp. 1 – 6). London, UK: British Computer Society.

Lorenz, A., & Oppermann, R. (2009). Mobile health monitoring for the elderly: Designing for diversity. *Pervasive and Mobile Computing*, *5*(5), 478-495.

Mandal, S., Kang, Z., & Millan, A. (2016). Data Visualization of Complex Eye Movements Using Directed Weighted Networks: A Case Study on a Multi-Element Target Tracking Task. In *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* (pp. 106-110). Los Angeles, USA: HFES.

Manzari, L., & Trinidad-Christensen, J. (2006). User-centered design of a web site for library and information science students: Heuristic evaluation and usability testing. *Information technology and libraries*, *25*(3), 163-169.

Mao, Y., Vredenburg, K., Smith, W., & Carey, T. (2005). The state of user-centered design practice. *Communications of the ACM*, *48*(3), 105-109.

Mintmire, E., Ling, C., & Shehab, R. (2013). Usability evaluation of a paratransit management system. In *Proceedings of the 2013 Industrial and Systems Engineering Research Conference* (pp. 2000-2009). San Juan, Puerto Rico: Institute of Industrial and Systems Engineers (IISE).

Monares, Á., Ochoa, S., Pino, J., Herskovic, V., Rodriguez-Covili, J., & Neyem, A. (2011). Mobile computing in urban emergency situations: Improving the support to firefighters in the field. *Expert systems with applications*, *38*(2), 1255-1267.

Morey, A., Barg-Walkow, H., & Rogers, A. (2017). Managing heart failure on the go: Usability issues with mHealth apps for older adults. In *Proceedings of the Human Factors and Ergonomics Society 61st Annual Meeting* (pp. 1-5). Los Angeles, CA: SAGE Publications.

National Weather Service. (2018). Natural Hazard Statistics: Weather Fatalities. Retrieved from http://www.nws.noaa.gov/om/hazstats.shtml.

Nayebi, F., Desharnais, J. M., & Abran, A. (2012). The state of the art of mobile application usability evaluation. In *Proceedings of the 25th Annual IEEE Canadian Conference on Electrical and Computer Engineering* (pp. 1-4), Canada: IEEE.

Nielsen, J. (1994a). *Usability engineering*. Mountain View, CA: Elsevier.

Nielsen, J. (1994b). Enhancing the explanatory power of usability heuristics. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Celebrating Interdependence* (pp. 152– 158). New York, USA: Association for Computing Machinery.

Nielsen, J. (1995a). 10 usability heuristics for user interface design. Retrieved from https://www.nngroup.com/articles/ten-usability-heuristics/.

Nielsen, J. (1995b). Usability inspection methods. In *Conference companion on Human factors in computing systems* (pp. 377-378). New York, USA: Association for Computing Machinery.

Nielsen, J. (1995c). Severity ratings for usability problems. Retrieved from https://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/.

Nielsen, J. (2009). Powers of 10: Time Scales in User Experience. Retrieved from https://www. nngroup.com/articles/powers-of-10-timescales-in-ux/.

Nielsen, J., & Molich, R. (1989). Teaching user interface design based on usability engineering. *Association for Computing Machinery SIGCHI Bulletin, 21*(1), 45–48.

Nielsen, J., & Molich, R. (1990). Heuristic evaluation of user interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems: Empowering People* (pp. 249–256). New York, USA: Association for Computing Machinery.

Nielsen. J. (1997). The Use and Misuse of Focus Groups. Retrieved from https://www.nngroup.com/articles/focus-groups/.

Nielsen. J. (2012). Usability 101: Introduction to Usability. Retrieved from https://www.nngroup.com/articles/usability-101-introduction-to-usability/.

Norman, D. (1988). *The design of everyday things*. New York: Doubleday.

Norman, D. (2013). *The design of everyday things: Revised and expanded edition*. New York, NY: Basic Books.

Norman, D., & Draper, S. (1986). *User Centered System Design: New Perspectives on Human-Computer Interaction*. New York, USA: CRC Press.

Noton, D., & Stark, L. (1971a). Eye movements and visual perception. *Scientific American*, *224*(6), 34–43.

Noton, D., & Stark, L. (1971b). Scanpaths in saccadic eye movements while viewing and recognizing patterns. *Vision Research*, *11*(9), 929–942.

Pak, R., & McLaughlin, A. (2010). *Designing displays for older adults*. Boca Raton, Fl: CRC Press.

Pan, D., Chen, N., & Rau, P. (2013). The acceptance and adoption of smartphone use among Chinese college students. In *Proceedings of the International Conference on Cross-Cultural Design* (pp. 450-458). Berlin, Heidelberg: Springer.

Panach, J., Condori-Fernández, N., Valverde, F., Aquino, N., & Pastor, Ó. (2008). Understandability measurement in an early usability evaluation for model-driven development: an empirical study. In *Proceedings of the 2nd ACM-IEEE international symposium on Empirical software engineering and measurement* (pp. 354-356). New York, USA: Association for Computing Machinery.

Poole, A., & Ball, L. (2006). Eye tracking in HCI and usability research. In *Proceedings of Encyclopedia of human computer interaction* (pp. 211-219). London, UK: IGI Global.

Poole, A., Ball, L., & Phillips, P. (2005). In search of salience: A response-time and eye-movement analysis of bookmark recognition. In *Proceedings of People and computers XVIII—Design for life* (pp. 363-378). London, UK: Springer.

Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S., & Carey, T. (1994). *Human-computer interaction*. Essex, UK: Addison-Wesley Longman Limited.

Pretorius, M., Calitz, A., & van Greunen, D. (2005). The added value of eye tracking in the usability evaluation of a network management tool. In *Proceedings of the 2005 annual research conference of the South African institute of computer scientists and information technologists on IT research in developing countries* (pp. 1-10). South Africa: South African Institute for Computer Scientists and Information Technologists.

Rogers, Y., Sharp, H., & Preece, J. (2011). *Interaction design: beyond human-computer interaction.* London, UK: John Wiley & Sons.

Saldana, J. (2015). *The Coding Manual for Qualitative Researchers*. Thousand Oaks, CA: Sage Publications.

Saleh, A., & Ismail, R. (2015). Usability evaluation frameworks of mobile application: a mini-systematic literature review. In *Proceedings of the 3rd Global Summit on Education* (pp. 231–239). Kuala Lumpur, MALAYSIA: GSE.

Sauro, J., & Dumas, S. (2009). Comparison of three one-question, post-task usability questionnaires. In *Proceedings of the Special Interest Group on Computer-Human Interaction conference on human factors in computing systems 27th Annual Meeting* (pp. 1599-1608). New York, NY: Association for Computing Machinery.

Schmiedl, G., Seidl, M., & Temper, K. (2009). Mobile phone web browsing: a study on usage and usability of the mobile web. In *Proceedings of the 11th International Conference on Human-Computer Interaction with Mobile Devices and Services* (pp. 245-256). Bonn, Germany: ACM.

Schnall, R., Rojas, M., Bakken, S., Brown, W., Carballo-Dieguez, A., Carry, M., & Travers, J. (2016). A user-centered model for designing consumer mobile health (mHealth) applications (apps). *Journal of biomedical informatics*, *60*(2), 243-251.

Schneiderman, B. (1992). *Designing the User Interface: Strategies for Effective Human-Computer Interaction* (2nd ed.). Reading, MA: Addison-Wesley.

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

Segall, N., Saville, J., L'Engle, P., Carlson, B., Wright, M., Schulman, K., & Tcheng, J. (2011). Usability evaluation of a personal health record. In *Proceedings of the American Medical Informatics Association annual symposium* (pp. 1233-1242). Washington, D.C.: American Medical Informatics Association.

Sekuler, B. & Ball, K. (1986). Visual localization: Age and practice. *Journal of the Optical Society of America*, *3*(6), 864-867.

Shackel, B. (1991). *Human factors for informatics usability*. New York, NY: Cambridge University Press.

Sharma, S., SIngh, P., Sharma, R., & Mahajan, A. (2012). Age based user interface in mobile operating system. *International Journal of Computer Science, Engineering and Applications*, *2*(2), 177-184.

Shrestha, S. (2007). Mobile web browsing: usability study. In *Proceedings of the 4th* *international conference on mobile technology, applications, and systems and the 1st international symposium on Computer human interaction in mobile technology* (pp. 187-194). Singapore: ACM.

Silva, P., Holden, K., & Nii, A. (2014). Smartphones, smart seniors, but not-so-smart apps: A heuristic evaluation of fitness apps. In *Proceedings of the International Conference on Augmented Cognition* (pp. 347-358). Cham, Switzerland: Springer.

Singhal, B. (2011). *Usability testing of iPhone* (Master's thesis). Retrieved from http://sdsu-dspace.calstate.edu/bitstream/handle/10211.10/1197/Singhal_Bikash.pdf?sequence=1.

Smith, L., & Chaparro, S. (2015). Smartphone text input method performance, usability, and preference with younger and older adults. *Human Factors: The Journal of Human factors and Ergonomics Society*, *57*(6), 1015-1028.

Stacy, A. (2017). Current Trends and Future Prospects of The Mobile App Market. Retrieved from https://www.smashingmagazine.com/2017/02/current-trends-future-prospects-mobile-app-market/.

Statista. (2018a). Market reach of the most popular mobile app categories in the United States as of February 2018. The Statistics Portal. Retrieved from https://www.statista.com/statistics/579302/top-app-categories-usa-reach/.

Statista. (2018b). Share of adults in the United States who owned a smartphone from 2015 to 2018, by age group. Retrieved from https://www.statista.com/statistics/489255/percentage-of-us-smartphone-owners-by-age-group/.

Statista. (2018c). Number of smartphone users in the United States from 2010 to 2022. Retrieved from https://www.statista.com/statistics/201182/forecast-of-smartphone-users-in-the-us/.

Sweller, J. (1988). Cognitive load during problem solving: effects on learning. *Cognitive Science*, *12*(2), 257–285.

Tangsoc, J. & Amelia, S. (2009). Usability of Wireless Application Protocol Interface in the Philippines. *Industrial Engineering & Management Systems*, *8*(1), 54-65.

Vredenburg, K., Mao, J., Smith, P., & Carey, T. (2002). A survey of user-centered design practice. In *Proceedings of the Special Interest Group on Computer-Human Interaction conference on Human factors in computing systems* (pp. 471-478). New York, NY: Association for Computing Machinery.

Wahl, W., & Tesch-Römer, C. (2001). Aging, sensory loss, and social functioning. In *N. Charness, D. C. Parks, & B. A. Sabel (Eds.), Communication, technology and aging: Opportunities and challenges for the future* (pp. 108-126). New York, NY: Springer.

Wahrendorf, M., Reinhardt, D., & Siegrist, J. (2013). Relationships of disability with age among adults aged 50 to 85: evidence from the United States, England and continental Europe. Retrieved from https: //journals.plos.org/plosone/article?id =10.1371/journal.pone.0071893.

Weather Decision Technology. (2018). Retrieved from http://wdtinc.com/.

Weather Radio (2018). Retrieved from http://weatherradioapp.com/.

Whitenton, W. (2016). Website Forms Usability: Top 10 Recommendations. Retrieved from https://www.nngroup.com/articles/web-form-design/.

Williams, J. (2004). *Developing performance support for computer systems: a strategy for maximizing usability and learnability*. Boca Raton, FL: CRC Press.

Witteman, H. O., Dansokho, S. C., Colquhoun, H., Coulter, A., Dugas, M., Fagerlin, A., ... & Ivers, N. (2015). User-centered design and the development of patient decision aids: protocol for a systematic review. *Systematic reviews*, *4*(1), 11.

Xu, Q., Erman, J., Gerber, A., Mao, Z., Pang, J., & Venkataraman, S. (2011). Identifying diverse usage behaviors of smartphone apps. In *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference* (pp. 329-344). New York, NY: Association for Computing Machinery.

Yen, S., & Esgin, E. (2015). Usability Evaluation of Web Based Educational Multimedia by Eye-Tracking Technique. *International Journal of Social Sciences and Education*, *5*(4), 590 – 603.

Zabini, F. (2016). Mobile weather apps or the illusion of certainty. *Meteorological Applications*, *23*(4), 663-670.

Appendix A: Exit Survey Questionnaire Used in Chapter 3

(1) What is your gender? M F

(2) What is your age?

(3) In your opinion, what are the top 3 features used on the Weather Radio application you just interacted with?

(4) What features do you feel are difficult to use? (List most difficult to least difficult).

(5) What features do you feel are easy to use? (List easiest first).

(6) How would you rate the overall usability of the Weather Radio application?

1 2 3 4 5

(1=very poor 2= poor 3=fair 4=good 5=very good)

(7) Please give any general comments or suggestions on the usability of weather apps:

Appendix B (1): Original and proposed STW messages Used in Chapter3

Original NWS STW message	Proposed STW message
WATCH COUNTY NOTIFICATION FOR WATCH 58 NATIONAL WEATHER SERVICE NORMAN OK 150 PM CDT WED MAR 30 2016 OKC015-017-019-027-031-033-047-049-051-053-067-071-073-081-083- 085-087-099-103-109-119-125-137-TXC009-023-077-485-310200-/O.NEW.KOUN.SV.A.0058.160330T1850Z-160331T0200Z/ THE NATIONAL WEATHER SERVICE HAS ISSUED SEVERE THUNDERSTORM WATCH 58 IN EFFECT UNTIL 9 PM CDT THIS EVENING FOR THE FOLLOWING AREAS IN OKLAHOMA THIS WATCH INCLUDES 23 COUNTIES IN CENTRAL OKLAHOMA CANADIAN…CLEVELAND…GRADY KINGFISHER…LINCOLN…LOGAN MCCLAIN…OKLAHOMA…PAYNE POTTAWATOMIE IN NORTHERN OKLAHOMA GARFIELD…GRANT…KAY…NOBLE IN SOUTHERN OKLAHOMA CARTER…GARVIN…JEFFERSON...LOVE…MURRAY… STEPHENS IN SOUTHWEST OKLAHOMA CADDO…COMANCHE…COTTON IN TEXAS THIS WATCH INCLUDES 4 COUNTIES IN NORTHERN TEXAS ARCHER…BAYLOR…CLAY THIS INCLUDES THE CITIES OF...ANADARKO… ARCHER CITY...ARDMORE... BLACKWELL...BLANCHARD...CHANDLER... CHICKASHA...CONCHO... DAVENPORT...DAVIS... DUNCAN...EL RENO...ENID...GUTHRIE... HENNESSEY...HENRIETTA...HINTON...HOLLIDAY... KINGFISHER...LAKESIDE CITY...LAMONT...LAWTON… LINDSAY...MARIETTA...MEDFORD... MEEKER… MOORE...MUSTANG...NEWCASTLE...NORMAN… OKARCHE...OKLAHOMA CITY...PAULS VALLEY… .PERRY...PONCA CITY...POND CREEK...PRAGUE… PURCELL...RINGLING...RYAN...SCOTLAND...SEYMOUR... SHAWNEE...SHEPPARDAFB...STILLWATER...STROUD... SULPHUR...TEMPLE...THACKERVILLE...TUTTLE... WAKITA...WALTERS...WAURIKA...WELLSTON... WICHITA FALLS...WYNNEWOOD AND YUKON.	Severe Thunderstorm Watch # 58 in 2016: For Counties of Oklahoma, Counties of Texas, and Cities that include the impacted counties. Time: 1:50 PM, Time Zone: Central Daylight Time (CDT), Day: Wednesday, Date: 03/30/2016 *Severe Thunderstorm Watch # 58 is in effect until 9 PM this evening for the following areas:* Counties under the Severe Thunderstorm Watch in Oklahoma: In Central Oklahoma: Cleveland – Grady - Canadian - Kingfisher – Lincoln - Logan - McClain Oklahoma - Payne – Pottawatomie In Northern Oklahoma: Kay - Garfield - Grant – Noble In Southern Oklahoma: Carter – Jefferson - Garvin - Love - Murray In Southwest Oklahoma: Comanche - Cotton – Caddo Counties under the Severe Thunderstorm Watch in Texas: In Northern Texas: Archer- Baylor- Clay- Wichita Cities Under the Severe Thunderstorm Watch in Oklahoma: Anadarko – Ardmore – Blackwell – Blanchard – Chandler – Chickasha – Concho – Davenport – Davis – Duncan – El Reno – Enid – Guthrie – Hennessey – Hinton – Henrietta – Holliday – Kingfisher – Lamont – Lawton – Lindsey –Medford – Meeker – Moore – Mustang – Newcastle – Norman – Okarche – Oklahoma City – Pauls Valley – Perry – Ponca City – Pond Creek – Prague – Purcell – Ringling – Ryan – Scotland – Seymour – Shawnee – Sheppard Afb – Stillwater – Stroud – Sulphur – Temple – Thackerville – Tuttle – Wakita – Walters – Waurika – Wellston – Wynnewood – Yukon. Cities Under the Severe Thunderstorm Watch in Texas: Archer City- Lakeside City – Marietta - Wichita Falls. *National Weather Service Center, Norman, OK*

Appendix B (2): Original and proposed WA messages Used in Chapter3

Original NWS WA message	Proposed WA message
URGENT - WEATHER MESSAGE...UPDATED NATIONAL WEATHER SERVICE NORMAN OK 456 PM CDT MON MAR 21 2016 OKZ007-008-011>013-015>020-022>031-035>040-044-221200- /O.CON.KOUN.WI.Y.0005.160322T1600Z-160323T0200Z/ GRANT-KAY-MAJOR-GARFIELD-NOBLE-DEWEY-CUSTER-BLAINE-KINGFISHER- LOGAN-PAYNE-WASHITA-CADDO-CANADIAN-OKLAHOMA-LINCOLN-GRADY-MCCLAIN- CLEVELAND-POTTAWATOMIE-SEMINOLE-KIOWA-JACKSON-TILLMAN-COMANCHE-STEPHENS-GARVIN-COTTON- INCLUDING THE CITIES OF...MEDFORD...POND CREEK... LAMONT... WAKITA...PONCA CITY...BLACKWELL...FAIRVIEW... ENID...PERRY... SEILING...VICI...TALOGA...LEEDEY... WEATHERFORD...CLINTON...WATONGA...GEARY... KINGFISHER...HENNESSEY...OKARCHE...GUTHRIE...STILLWATER... CORDELL...BURNS FLAT...SENTINEL... ANADARKO...HINTON... YUKON...CONCHO...EL RENO...MUSTANG... OKLAHOMA CITY...CHANDLER...STROUD...PRAGUE...MEEKER...DAVENPORT... WELLSTON...CHICKASHA...TUTTLE...PURCELL...NEWCASTLE... BLANCHARD...NORMAN...MOORE...SHAWNEE...SEMINOLE... WEWOKA...HOBART...SNYDER... ALTUS...FREDERICK...LAWTON... DUNCAN...PAULS VALLEY...LINDSAY...WYNNEWOOD... WALTERS...TEMPLE 456 PM CDT MON MAR 21 2016 ...WIND ADVISORY REMAINS IN EFFECT FROM 11 AM TO 9 PM CDT TUESDAY... * TIMING...11 AM TO 9 PM. * WINDS...SOUTH TO SOUTHWEST 25 TO 35 MPH WITH GUSTS 40 TO 50 MPH. * IMPACTS...DRIVING COULD BECOME DIFFICULT ESPECIALLY IN HIGH PROFILE VEHICLES. ANY LOOSE OUTDOOR ITEMS COULD ALSO BLOW AROUND. PRECAUTIONARY/PREPAREDNESS ACTIONS... BE CAREFUL IF YOU HAVE TO TRAVEL OR IF YOU ARE WORKING OR PLAYING OUTSIDE.	Urgent – Weather Message: Updated Wind Advisory: For Counties of Oklahoma. Time: 4:56 PM, Time Zone: Central Daylight Time (CDT), Day: Monday, Date: 03/21/2016 Wind Advisory remains in effect from 11 AM CDT on Monday to 9 PM CDT on Tuesday Impacts: Driving could become difficult especially in tall vehicles. Any loose outdoor items could also blow around. Avoid riding motorcycles or bicycles Precautionary/Preparedness Actions: Be careful if you have to travel or if you are working or playing outside. Counties: OK: Grant- Kay- Major- Garfield- Noble- Dewey- Custer- Blaine- Kingfisher- Logan- Payne- Washita- Caddo-Canadian- Oklahoma- Lincoln- Grady- McClain- Cleveland- Pottawatomie- Seminole- Kiowa- Jackson-Tillman- Comanche- Stephens- Garvin- Cotton. Cities: OK: Medford- Pond Creek – Lamont- Wakita- Ponca City- Blackwell- Fairview- Enid- Perry- Seiling- Vici- Taloga- Leedey- Weatherford- Clinton- Watonga- Geary- Okeene- Kingfisher- Hennessey- Okarche- Guthrie- Stillwater- Cordell- Burns Flat- Sentinel- Anadarko- Hinton- Yukon- Concho- El Reno- Mustang- Oklahoma City- Chandler- Stroud- Prague- Meeker- Davenport- Wellston- Chickasha- Tuttle- Purcell- Newcastle- Blanchard- Norman- Moore- Shawnee- Seminole- Wewoka- Hobart- Snyder- Altus- Frederick- Lawton- Duncan- Pauls Valley- Lindsay- Wynnewood- Walters- Temple. *National Weather Service Center, Norman, OK*

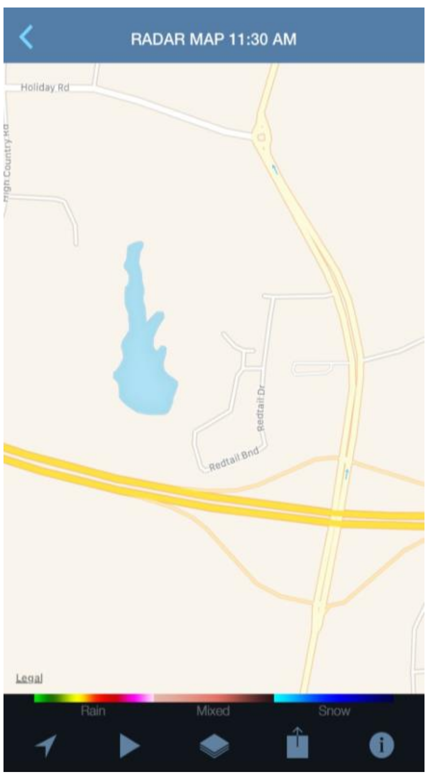
Appendix C: Focus Group Questions Used in Chapter 4

(1) Apart from accurate weather data, what influences your decision to download a particular mobile weather app?

(2) What are your top 3 features (e.g. location search, alert messages, weather forecasts, weather maps, settings customization…etc.) on mobile weather apps? Why?

(3) Please tell me about your positive and negative experiences in using mobile weather apps.

(4) In the screenshot shown below, which is a map feature on a particular weather app, how do you perceive the use of the icons at the bottom? Can your experience be better with such icons? If yes, How so?



(5) What is your perception about using an exact location (e.g. house address or school name) vs. adding a nearby city or postal code, in terms of obtaining reliable weather forecasts?

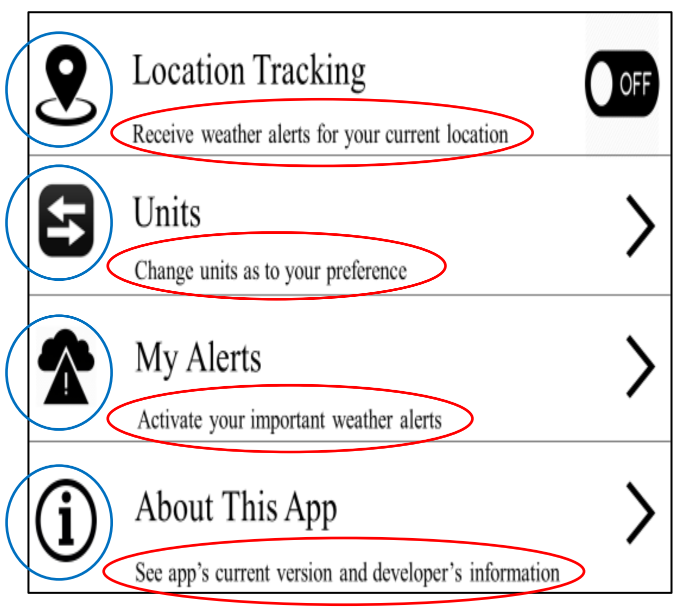
For those who go with using exact location, which of the following would be easier to you:

A) Searching the app's map for your exact location. Why?

B) Typing the location name/address in the search bar with effective search index (ex. Auto suggestions). Why?

(6) What do you think about adding descriptive information (e.g. circled information in red) and icons (e.g. circled in blue) to apps' features?

Would your answer change based on the size of the smartphone on which the weather app being shown? Why?



(7) Below, a sample of weather alert messages pushed to an app's users during a severe weather condition, please tell me what you think about it (e.g. in terms of message comprehension, message length, text style, scope of message...etc.) If you think alert messages can be improved, how so? Consider time-critical weather conditions (e.g. tornado warning) in your answer.

"WATCH COUNTY NOTIFICATION FOR WATCH 58
NATIONAL WEATHER SERVICE NORMAN OK
150 PM CDT WED MAR 30 2016
OKC 015-017-019-027-031-033-047-049-051-053-067-071-073-081-083- 085-087-099-103-109-119-125-137-TXC009-023-077-485-310200-/O.NEW.KOUN.SV.A.0058.160330T1850Z-160331T0200Z/
THE NATIONAL WEATHER SERVICE HAS ISSUED SEVERE THUNDERSTORM WATCH 58 IN EFFECT UNTIL 9 PM CDT THIS EVENING FOR THE FOLLOWING AREAS
IN OKLAHOMA THIS WATCH INCLUDES 23 COUNTIES
IN CENTRAL OKLAHOMA
CANADIAN CLEVELAND GRADY KINGFISHER LINCOLN LOGAN
MCCLAIN OKLAHOMA PAYNE POTTAWATOMIE
IN NORTHERN OKLAHOMA
GARFIELD GRANT KAY NOBLE
IN SOUTHERN OKLAHOMA
CARTER GARVIN JEFFERSON LOVE MURRAY
STEPHENS
IN SOUTHWEST OKLAHOMA
CADDO COMANCHE COTTON
IN TEXAS THIS WATCH INCLUDES 4 COUNTIES
IN NORTHERN TEXAS
ARCHER BAYLOR CLAY WICHITA
THIS INCLUDES THE CITIES OF...ANADARKO...ARCHER CITY...ARDMORE...BLACKWELL...BLANCHARD...CHANDLER...CHICKASHA...CONCHO...DAVENPORT...DAVIS...DUNCAN...EL RENO...ENID...GUTHRIE...HENNESSEY... HENRIETTA...HINTON...HOLLIDAY...KINGFISHER...LAKESIDECITY...LAMONT...LAWTON...LINDSAY...MARIETTA...MEDFORD...MEEKER...MOORE...MUSTANG...NEWCASTLE...NORMAN...OKLAHOMA CITY...PAULS VALLEY...
PERRY...PONCA CITY...POND CREEK...PRAGUE...PURCELL...RINGLING...RYAN...SCOTLAND...SEYMOUR...
SHAWNEE...SHEPPARD AFB...STILLWATER...STROUD...SULPHUR...TEMPLE...THACKERVILLE...TUTTLE...WAKITA...

WALTERS...WAURIKA...WELLSTON...WICHITA FALLS...WYNNEWOOD AND YUKON."

(8) Some users like apps that enable them to access and control (turn on/off) alerts and sub-alerts for all types of weather. During severe weather conditions, the users would receive alert messages for their turned-on alerts and sub-alerts.

On the other hand, some users prefer apps that mainly show the critical weather alerts. During severe weather conditions, the users would receive alert messages for their turned-on time-critical alerts (e.g. warning alerts). Non-time-critical alerts (e.g. wind watch) are not automatically sent to users as messages; those alerts would only appear with a representative symbol on the screen with the impacted location.

Which type of apps do you prefer? Why?

(9) Of all the things we have discussed, which is the most important to you? Why? Do you have any further comments about the usability of the weather apps you have interacted with or any suggestions that may enhance your future weather apps experience?