

Usability Evaluation of Mobile Weather Hazard Alert Applications

Abdulrahman Khamaj^{1,2} and Ziho Kang¹

¹School of Industrial and Systems Engineering, University of Oklahoma
Oklahoma, United States

²School of Industrial Engineering, Jazan University
Jazan, Saudi Arabia

Corresponding author's Email: khamaj@ou.edu

Abstract: Cell phones enable us to receive and respond to critical incidents, such as: severe storms, tornadoes, and flash floods. However, due to the small display size of cell phones, and regardless of simplified symbols or alert messages, it is possible to overlook users' ability to interact with the available features and understand the messages in a timely manner. Untrained and trained users of the Weather Radio application participated in an experiment to perform three search tasks; (task 1: location search, task 2: alert settings, and task 3: map settings). In task 4, they evaluated two types of weather alert messages: original National Weather Service (NWS) messages vs. filtered (proposed) messages. By recording users' completion time on the search tasks, the results showed that the time of the typing in text bar method for task 1 was significantly less than the pin on map method, while much more time was required to complete tasks 2 and 3 by the untrained users compared to the trained users. It was also revealed that the proposed messages were more effective than the original messages by both user groups. This research of user-centered designs provides a foundation to support the designs of time-critical mobile alert systems.

Keywords: Usability Evaluation, Mobile, Weather application.

1. Introduction

Numerous smartphone (mobile) weather alert applications are developed and made available for users. In addition, large numbers of users heavily rely on accessing weather information and critical weather-related events through the applications embedded in their mobile devices [see 1.1 for statistics]. However, many of these application interfaces may not be highly usable and properly designed. Furthermore, the research regarding the usability of these applications is still significantly lacking. This study aims to: evaluate the usability of mobile weather alert applications, determine the issues that result from users' interaction with the applications, and propose and test alternative approaches. Findings from this study could enhance the currently lacking research area of mobile weather applications as well as help application developers to consider the results and recommendations for better future designs.

1.1. Mobile Weather Alert 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 National Weather Service (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.

There are several sources that deliver weather information to the public through televisions, radios, and most commonly used, smartphones. Recent statistics showed that more than 2 billion people globally used smartphones in 2016, with around 207.2 million users during the same time in the United States; the number is estimated to increase rapidly ("Statistica," 2017). Smartphones tend to make lives easier as users can get full and easy access to the technology through their devices while they are on the move (Nayebi et al., 2012).

The increasing number of smartphone users have encouraged companies and technology experts to develop vast numbers of applications to be used by mobile consumers. For instance, more than 150,000 applications are available for Android users and around 350,000 users activate applications daily (Xu et al., 2011). With the easy accessibility of information shown on mobile devices, weather applications are becoming increasingly popular. For example, over 5.2 million users have installed the Weather Radio application created by Weather Decision Technologies (WDT) on their devices ("Weather Decision

Technology," 2016).

However, these applications significantly vary in quality as many application 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. Ignoring the usability of these applications, and not thoroughly considering the different characteristics of the ultimate users, may lead to severe consequences. With the growing dependency of people on weather information presented through mobile weather applications, poorly designed applications may fail to convey the weather alerts (the risk level associated with the weather feature) properly; especially during severe weather situations, such as tornados, that require appropriate reaction in a timely manner.

1.2. Usability Evaluation Methods of Mobile Applications

Several definitions of usability are available in the literature. However, 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).

Usability evaluation is considered one of the most important techniques to test the quality and discover the challenges and limitations within smartphone applications (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).

There are several usability evaluation methods considered in the literature. Observational method (Khanum & Trivedi, 2012), focus group method (Krueger & Casey, 2002), GOMS method (John & Kieras, 1996), and recently eye tracking method (Poole & Ball 2006) are among the most popular and frequently used usability methods. In particular, the observational method includes the investigation of attributes such as: efficiency, learnability, and satisfaction which are advocated by numerous usability researchers. For example, The International Organization for Standardization (ISO 9241-11) (1997) identified effectiveness, efficiency, and satisfaction as the measures that determine whether or not a system is usable. Nielsen (1994), proposed three additional important attributes to what ISO 9241-11 (1997) identified: learnability, memorability and errors.

Among the usability attributes listed above, the two important attributes of our interest are: efficiency of use and user satisfaction. Efficiency of use refers to how fast and accurately a user can accomplish a task (Frekjm et al., 2000) and is usually measured through the time it takes for a user to complete a task. Short completion time indicates that the system is simple (easy to interact with), easy to learn, easy to remember, and error free or with less errors. User satisfaction refers to how a user feels overall about the interaction with a particular system (Han et al., 2004) and it is gauged subjectively (e.g. through Likert rating scale). Customer (user) satisfaction is one of the most important factors for the success of any product or system and helps convey overall opinions or feedback on the product or system of interest.

1.3. Current Usability Research of Mobile Weather Alert Applications

Several studies have focused on examining the usability of mobile applications in different specific areas such as in tourism (Geven et al., 2006; Shrestha 2007; Ahmadi and Kong 2008; Schmiedl et al., 2009) and geography (van Elzakker et al., 2008). However, very few studies were directly pertaining to weather alert applications. Throughout the evaluation on this area, two consecutive studies by Singhal (2011) and then Alluri (2012) examined the usability issues in the interface design of all the originally built-in mobile applications 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 applications in general. Both studies revealed multiple issues in each application such as lack of visibility, lack of affordance, and poor consistency. For example, the lack of visibility was present in the weather application in both studies, where participants could not easily see the weather information icon "i" because of its very small size.

The most applicable usability evaluation study of mobile weather alert applications to this study 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 mobile "Weather Channel" application 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's et al. study (2015), 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 limited number of users (3-5 users), from a statistical standpoint, at least twenty participants should be involved (Nielsen, 2012); Drogalis's et al. (2015) study included only six participants.

Previous research has captured a basic understanding of the traditional usability attributes. This research has created the groundwork for other researchers to apply modern methods and metrics to evaluate the usability. Beyond the study done by Drogalis et al. (2015), there has still been little research done on mobile weather alert applications.

1.4. Specific Issues with the Existing Weather Applications

As most of the popular mobile weather applications include similar embedded features, four important features are outlined in this paper to determine their usability issues. In this paper, the widely used WDT's Weather Radio application is considered as the sample representative of these popular weather applications.

- (1) Location search: Many mobile applications provide a feature to navigate through the embedded Google map using the users' index finger, and to place (or relocate) a red pin shaped icon to indicate the point of interest. Although the process seems simple, this may not necessarily be true when the end-users actually interact with the feature within the small display. Specifically, the users need to press the pin icon using the finger and hold for more than one second to control (i.e. drag to a specific resident address on the map) the pin, to which many users may not be accustomed or even know how to use. In addition, if the users do not know the geographical layout of the point of interest, not only will it take time to pan and zoom within the Google map to find the point of interest, but also the users have to pan to the original pin location and drag it.
- (2) Alert settings: Several weather applications give users access to enable and/or disable some weather alert types. However, many weather alert applications may have an overwhelming number of alerts. This may hinder filtering out the most critical alerts properly and in a timely fashion. For example, the Weather Radio application includes twelve alerts such as flood, winter, wind, and thunderstorms and tornados; within each alert, there are several sub-alerts that the users can enable and/or disable. Specifically, the wind alert includes sixteen sub-alerts such as: high wind advisory, wind advisory, dust storm warning, and blowing dust advisory. Furthermore, with the small display size of mobile phones, application developers may use abbreviation and jargons. For example, in the Weather Radio application, to get access to the alert types, users need to first tap "NWS ALERTS" shown on the settings screen. NWS remains ambiguous for first time users; NWS stands for National Weather Services. But, because first time users may not know what "NWS" means, they may try each option in the settings list.
- (3) Map settings: Mobile weather applications usually include features to manipulate the settings of the map based on the user preference. In particular, the map settings in the weather alert applications usually enable users to choose one of several different map types such as standard, satellite, or hybrid. Even though these features provide flexibility in changing the map settings, getting access to the menu of these settings may not be easy. Specifically, due to the limited screen size, some mobile applications try to address this issue through implementing additional secondary menus. The location of these menus varies depending on the application developers' perspective. For example, the map settings menu in the Weather Radio application is placed in an invisible and counter-intuitive location in which users need to tap the map to be able to see the menu. Although this is one way to accommodate the various embedded features, first time users may get confused when searching for the required menu.
- (4) Weather alert messages: The Weather Radio application, as well as most of the popular mobile weather applications, includes a feature of showing weather alert messages that are sent by NWS or any National Oceanic and Atmospheric Administration's organizations. This feature gives users detailed information about potential weather hazards, possible impacts, and precautionary advices. However, these notifications are usually delivered to the end-user through the applications exactly as received from NWS devices; the applications automatically push the raw alert messages to users [see Figure 1.a]. Most of the information provided includes undefined codes, technical terms, and counter-intuitive, unorganized data. For example, the codes (OKC015, TXC009...) shown in Figure (1.b) represent the geographical codes for the areas under alert,

which may not be understood by general users. These issues may cause major difficulties to users in understanding the alert notifications and taking the required actions.



Figure 1.a. Current method of delivering weather alerts to the application’s end-user based on a “push” system

```
“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/”
```

Figure 1.b. Partial sample of a header information relayed on the Weather Radio application

As stated earlier, this study aims to measure both untrained and trained users’ performance on the given tasks through comparing their performance to each other. In addition, this paper aims to propose other approaches (i.e. applying contextual information to the weather alert messages) and test to what extent the proposed approaches help enhance the usability of the weather alert applications in comparison to the existing features. More specifically, this paper focuses on investigating four important features in the Weather Radio application: (1) searching for locations using two methods: dragging the pin on the map and typing the address in the application’s text bar; (2) changing the alert settings; (3) changing the map settings; and (4) comparing two sets of weather alert messages. In addition to the given tasks, an exit survey was given to participants asking about their experience during the experiment, as well as about the overall usability of the Weather Radio application.

2. Method

2.1. 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 smartphone 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 application (trained users) and 2) 20 first time users (untrained users). One of the study researchers provided the training sessions to the trained users group. The age of users ranged from 21 to 44 years (Mean (M) = 24.70, Standard Deviation (SD) = 4.89 years). Both the untrained and trained users performed all the given tasks. Even though the usability issues are mostly determined from first time users’ interaction with interfaces, trained 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 applications.

2.2. Apparatus

The Weather Radio (version: 3.0.5) (<http://weatherradioapp.com/>) was installed and run on a smartphone (iPhone 6) [see Figure 2]. A stopwatch was used to collect the response times for each of the given Weather Radio application tasks. The demographic survey, the different types of alert messages, and the exit survey were printed out on paper.

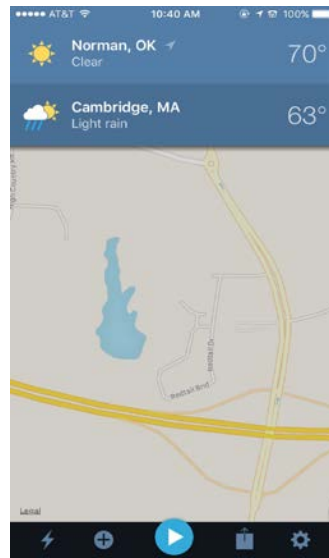


Figure 2. WDT's Weather Radio application on iPhone 6

2.3. 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 application's features. In addition, they were given time to practice navigating the application's interface by themselves and to ask questions if needed; they were asked to verbally state the following: "I am ready to begin the experiment" once they felt comfortable with the application. 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 application 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 mobile device, while 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 mobile device by recording the tasks' completion time.

The weather alert messages task was not accomplished on the Weather Radio application because the original weather alert messages only appear when there is a weather threat 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 sub-section 2.3.4]. 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 towards the overall usability of the application.

2.4. Tasks

2.4.1. Location Search Task

The location search task was to find a specific location using two approaches: pin icon allocation approach and the typing approach. The pin icon allocation approach was a feature implemented by the Weather Radio application. The purpose of this feature was to search the embedded Google map for a specific location for which a user needs 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 application, 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 the 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". To successfully accomplish this task with this approach, it was required to perform multiple steps: (1) click on the "+" icon, (2) type the city and the state in the task

bar that will pop up the Google map showing the city and a red pin icon randomly located near the center of the city, (3) pan and zoom the map to find Mount Auburn Hospital, (4) press and hold the pin until it lifts, and (5) move the pin icon into the Mount Auburn Hospital [see Figure 3 for visual explanation]. 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 application’s text bar.

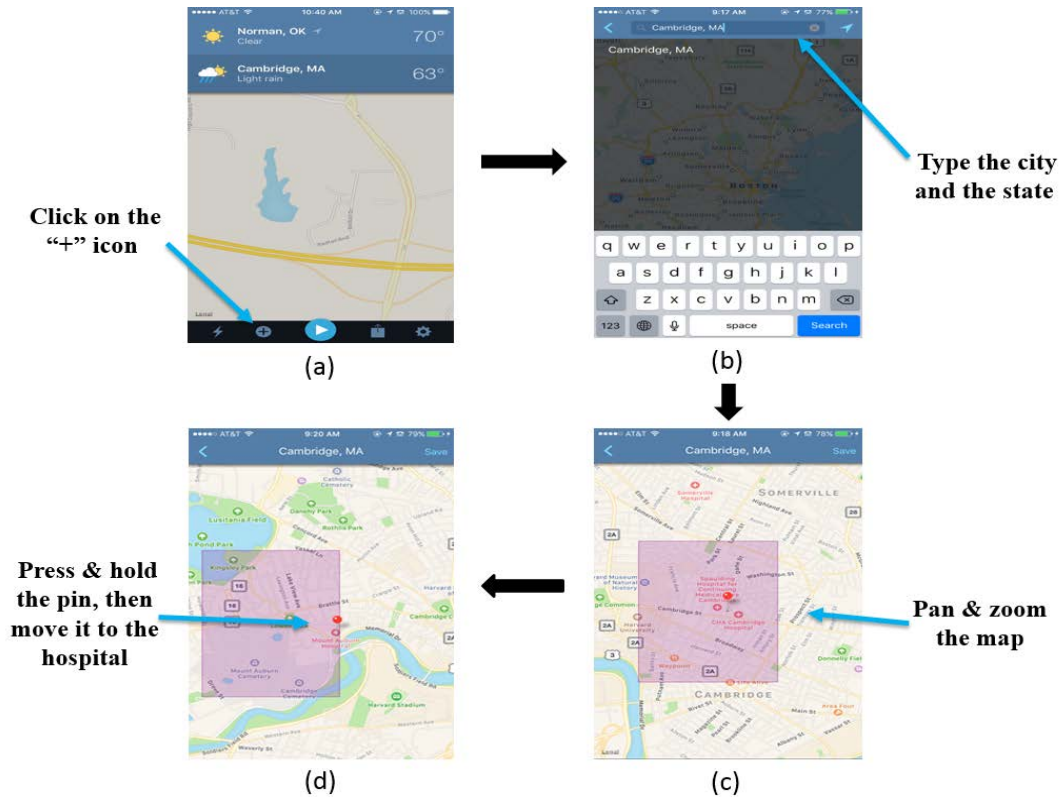


Figure 3. Process of locating the Mount Auburn Hospital using the pin icon

2.4.2. Alert Settings Task

The alert settings task was to change certain weather alert notifications. In particular, participants were asked to perform the following: “Please enable (turn on) Tornado Warning and Severe Thunderstorm Warning and disable (turn off) Tornado Watch and Severe Thunderstorm Watch”. To successfully accomplish this task, it was required to perform four steps: (1) click on the Gear icon from the home screen menu, (2) click on “NWS Alerts” from the setting options, (3) click on Thunderstorms and Tornadoes from NWS Alerts list, and (4) enable Tornado Warning and Severe Thunderstorm Warning and disable Tornado Watch and Severe Thunderstorm Watch. See Figure 4 for visual explanation.

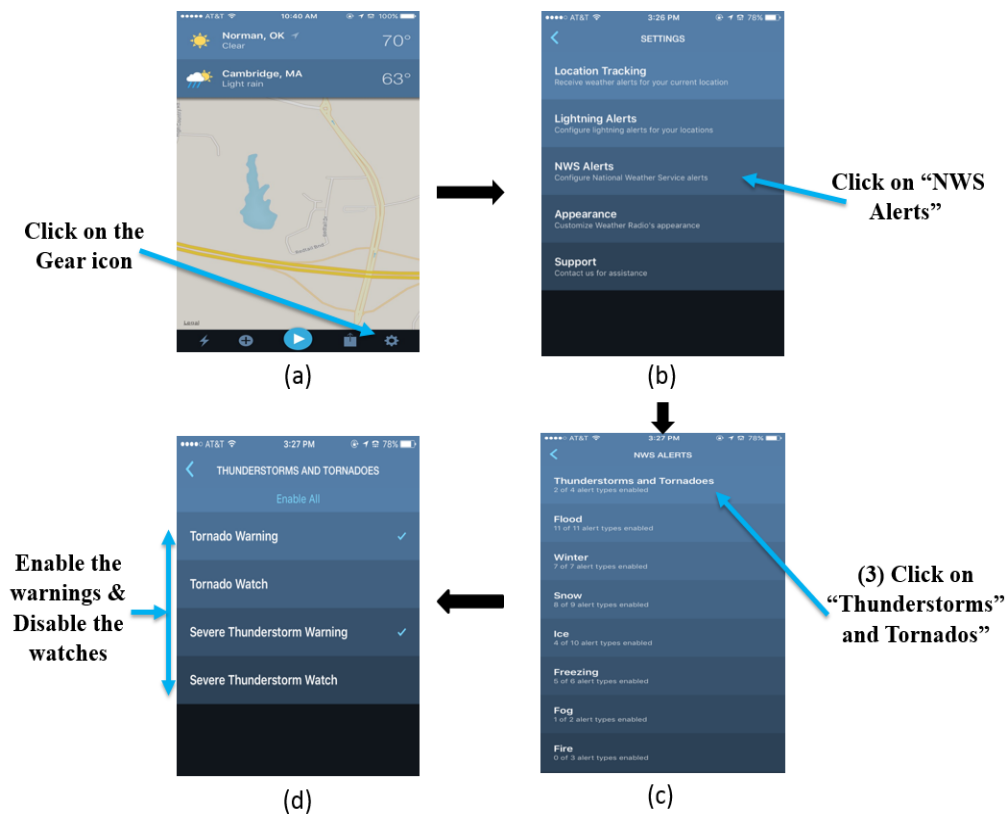


Figure 4. Process of adjusting alert settings

2.4.3. Map Settings Task

The map settings task was to change the settings of the map type and the weather layer. More specifically, participants were asked to perform the following: "Please change the map type from Standard to Hybrid and the weather layer from Radar to Clouds". The steps required for successfully accomplishing this task were: (1) tap the map on the home screen, (2) click on the information icon "i" from the pop up menu, and (3) from map setting options change the map type from Standard to Hybrid and the weather layer from Radar to Clouds. See Figure 5 for visual explanation.



Figure 5. Process of adjusting map settings

2.4.4. Weather Alert Message Evaluation Task

This task included two examples of weather alert messages that previously appeared on the Weather Radio application to alert users about current and future weather threats. The first weather message, Severe Thunderstorm Watch (STW) message, appeared on the application on March 30, 2016 to warn users about a severe thunderstorm watch and the second one, wind advisory (WA) message, 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 Table 1]. Similarly, the original version of the WA message was compared with the proposed version of the WA message [see Table 3]. 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, organized information 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.

2.4.4.1 STW Messages

Table 1. Original and Proposed STW messages

Original NWS STW message	Proposed STW message
<p>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/</p> <p>THE NATIONAL WEATHER SERVICE HAS ISSUED SEVERE THUNDERSTORM WATCH 58 IN EFFECT UNTIL 9 PM CDT THIS EVENING FOR THE FOLLOWING AREAS</p> <p>IN OKLAHOMA THIS WATCH INCLUDES 23 COUNTIES IN CENTRAL OKLAHOMA CANADIAN...CLEVELAND...GRADY KINGFISHER...LINCOLN...LOGAN MCCLAIN...OKLAHOMA...PAYNE POTTAWATOMIE</p> <p>IN NORTHERN OKLAHOMA GARFIELD...GRANT...KAY...NOBLE</p> <p>IN SOUTHERN OKLAHOMA CARTER...GARVIN...JEFFERSON...LOVE...MURRAY...STEPHENS</p> <p>IN SOUTHWEST OKLAHOMA CADDO...COMANCHE...COTTON</p> <p>IN TEXAS THIS WATCH INCLUDES 4 COUNTIES IN NORTHERN TEXAS ARCHER...BAYLOR...CLAY</p> <p>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.</p>	<p>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</p> <p><i>Severe Thunderstorm Watch # 58 is in effect until 9 PM this evening for the following areas:</i></p> <p>Counties under the Severe Thunderstorm Watch in Oklahoma:</p> <p>In Central Oklahoma: Cleveland – Grady - Canadian - Kingfisher – Lincoln - Logan - McClain Oklahoma - Payne – Pottawatomie</p> <p>In Northern Oklahoma: Kay - Garfield - Grant – Noble</p> <p>In Southern Oklahoma: Carter – Jefferson - Garvin - Love - Murray</p> <p>In Southwest Oklahoma: Comanche - Cotton – Caddo</p> <p>Counties under the Severe Thunderstorm Watch in Texas:</p> <p>In Northern Texas: Archer- Baylor- Clay- Wichita</p> <p>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.</p> <p>Cities Under the Severe Thunderstorm Watch in Texas: Archer City- Lakeside City – Marietta - Wichita Falls.</p> <p><i>National Weather Service Center, Norman, OK</i></p>

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

	Original NWS message statements	Proposed message statements
Header	<p>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./”</p>	<p>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.”</p>
Use of Delimiter	<p>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...ANADARKO...ARCHER CITY...ARDMORE...BLACKWELL...BLANCHARD...CHANDLER...CHICKASHA...CONCHO...DAVENPORT...DAVIS...DUNCAN...EL RENO...ENID...GUTHRIE...HENNESSEY...HENRIETTA...HINTON...HOLLIDAY...KINGFISHER.”</p>	<p>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.”</p>
Letters Format	<p>3) I find using only upper-case letters significantly enhanced the readability of this message.</p>	<p>3) I find using both upper-case and lower-case letters significantly enhanced the readability of this message.</p>
Satisfaction	<p>4) Overall, I am satisfied with the content and organization of this message.</p>	<p>4) Overall, I am satisfied with the content and organization of this message.</p>

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 2 for more details]

2.4.4.2 WA Messages

Table 3. Original and Proposed WA messages

Original NWS WA message	Proposed WA message
<p>URGENT - WEATHER MESSAGE...UPDATED</p> <p>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/</p> <p>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-</p> <p>INCLUDING THE CITIES OF...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</p> <p>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.</p> <p>* WINDS...SOUTH TO SOUTHWEST 25 TO 35 MPH WITH GUSTS 40 TO 50 MPH.</p> <p>* IMPACTS...DRIVING COULD BECOME DIFFICULT ESPECIALLY IN HIGH PROFILE VEHICLES. ANY LOOSE OUTDOOR ITEMS COULD ALSO BLOW AROUND.</p> <p>PRECAUTIONARY/PREPAREDNESS ACTIONS... BE CAREFUL IF YOU HAVE TO TRAVEL OR IF YOU ARE WORKING OR PLAYING OUTSIDE.</p>	<p>Urgent – Weather Message: Updated</p> <p>Wind Advisory: For Counties of Oklahoma. Time: 4:56 PM, Time Zone: Central Daylight Time (CDT), Day: Monday, Date: 03/21/2016</p> <p>Wind Advisory remains in effect from 11 AM CDT on Monday to 9 PM CDT on Tuesday</p> <p>Impacts: Driving could become difficult especially in tall vehicles. Any loose outdoor items could also blow around. Avoid riding motorcycles or bicycles</p> <p>Precautionary/Preparedness Actions: Be careful if you have to travel or if you are working or playing outside.</p> <p>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.</p> <p>Cities: OK: Medford- Pond Creek – Lamont- Wakita- Ponca City- Blackwell- Fairview- Enid- Perry- Seiling- Vici- Taloga- Leedeey- 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.</p> <p><i>National Weather Service Center, Norman, OK</i></p>

Table 4. Survey statements for the original and proposed wind advisory messages

	Original NWS message statements	Proposed message statements
Information Location	<p>1) I believe that the Wind Advisory information including the Impacts information located at the end of this message is appropriate.</p> <p>“...WIND ADVISORY REMAINS IN EFFECT FROM 11 AM TO 9 PM CDT TUESDAY...”</p> <p>* TIMING...11 AM TO 9 PM.</p> <p>* WINDS...SOUTH TO SOUTHWEST 25 TO 35 MPH WITH GUSTS 40 TO 50 MPH.</p> <p>* IMPACTS...DRIVING COULD BECOME DIFFICULT ESPECIALLY IN HIGH PROFILE VEHICLES. ANY LOOSE OUTDOOR ITEMS COULD ALSO BLOW AROUND.”</p>	<p>1) I believe that the Wind Advisory information including the Impacts information located at the top of this message is appropriate.</p> <p>“Wind Advisory: For Counties of Oklahoma.</p> <p>Time: 4:56 PM, Time Zone: Central Daylight Time (CDT), Day: Monday, Date: 03/21/2016</p> <p>Wind Advisory remains in effect from 11 AM CDT on Monday to 9 PM CDT on Tuesday.</p> <p>Impacts: Driving could become difficult especially in tall vehicles. Any loose outdoor items could also blow around. Avoid riding motorcycles or bicycles.”</p>
Word Expressions	<p>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.”</p>	<p>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.”</p>
Terminology	<p>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.”</p>	<p>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”</p>
Satisfaction	<p>4) Overall, I am satisfied with the content and organization of this message.</p>	<p>4) Overall, I am satisfied with the content and organization of this message.</p>

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 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 4 for more details].

2.5. 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: training and approach type. The levels of training were untrained and trained users. The levels of the approach type were pin on map and typing in 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: training with the same levels as in location search and alert messages tasks.

Two dependent variables were included in this study: task completion time and survey Likert rating score. The task completion time was used to assess the users’ performance on the location search, alert settings, and map settings 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 [see section 2.4.4. for details].

2.6. Data Analysis

Paired sample t-test and/or independent sample t-test was used for each of the experiment's tasks. Specifically, both the paired sample t-test and the independent t-test were used for the location search and alert messages tasks. The paired sample t-test was used to compare the two location search and alert messages approaches once by the untrained users and another by the trained users; the independent t-test was used to compare both the untrained and trained users' data on each approach for the location search and each comparison item for the alert messages. In addition, the independent sample t-test was used for both the map settings and the alert settings tasks to compare the data collected from the untrained users with the data collected from the trained users. Finally, for the exit user satisfaction survey, descriptive statistics analysis or qualitative content analysis was used.

3. Results

3.1. Weather Radio Application Tasks Results

3.1.1. Location Search Task

Figure 6 shows graph comparisons between the two approaches of the location search task as well as between the two user groups in terms of the mean completion time with Standard Error (SE) bars.

3.1.1.1 Location Search (Pin on Map) vs. (Text Bar): Untrained Users & Trained Users

A paired sample *t* test was conducted to compare the performance of untrained users on the location search task using the pin on the map versus using the text bar, in terms of the task completion time. The results showed that there was a significant difference between the two approaches, $t(19) = 4.72, p < .001$, indicating that the mean completion time of the location search using the application's text bar ($M = 51.45s, SD = 13.66s, N = 20$) was significantly less than the mean completion time for the location search using the pin on the map ($M = 150.40s, SD = 95.79s, N = 20$) when both were performed by untrained users.

In addition, the paired sample *t* test was performed for the location search task using the pin on the map versus using the text bar when accomplished by trained users. It was revealed that there was a statistically significant difference between the two approaches, $t(19) = 5.41, p < .001$, indicating that the mean completion time of the location search using the application's text bar ($M = 44.55s, SD = 9.52s, N = 20$) was significantly less than the mean completion time for the location search using the pin on the map ($M = 128.25s, SD = 68.02s, N = 20$) when both were performed by trained users.

3.1.1.2. Location Search (Pin on Map): Untrained Users vs. Trained Users

An independent Sample *t* test was conducted to compare the performance of untrained users to trained users on the location search using the pin on map approach. The results showed no significant difference between the untrained users ($M = 150.40s, SD = 95.79s, N = 20$) and the trained users ($M = 128.25s, SD = 68.02s, N = 20$), $t(38) = .843, p = .404$.

3.1.1.3. Location Search (App's Text Bar): Untrained Users vs. Trained Users

An independent Sample *t* test was conducted to compare the performance of untrained users to trained users on the location search using the typing approach. Similar to the comparison results in 3.1.1.2, it was revealed that there was no significant difference between the untrained users ($M = 51.45s, SD = 13.66s, N = 20$) and the trained users ($M = 44.55s, SD = 9.52s, N = 20$), $t(38) = 1.853, p = .072$.

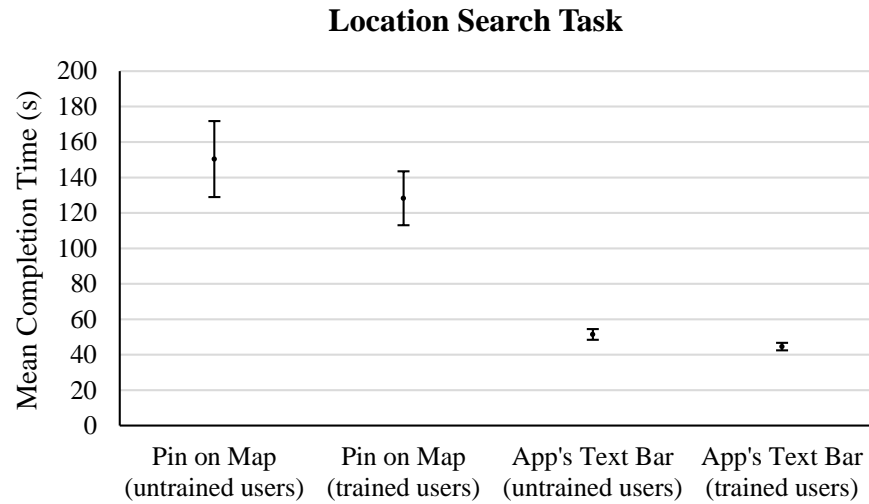


Figure 6. Plot of task completion time for the location search

3.1.2. Alert Settings: Untrained users vs. Trained Users

Figure 7 shows graph comparison between the untrained users and the trained users on the alert settings task in terms of the mean completion time with SE bars.

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

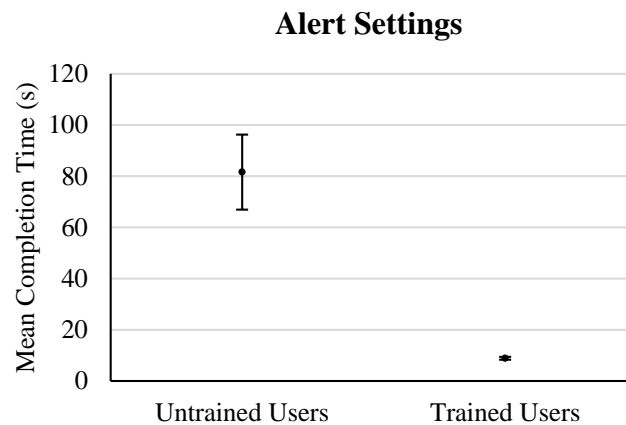


Figure 7. Plot of task completion time for the alert settings

3.1.3. Map Settings: Untrained users vs. Trained Users

Figure 8 shows graph comparison between the untrained users and the trained 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 untrained users and the trained users, in terms of the mean completion time on the map settings task. The results revealed a statistical significant difference, $t(38) = 8.459, p < .001$, indicating that the mean completion time of the trained users ($M = 3.95s, SD = .94s, N = 20$) was significantly less than the mean completion time of the untrained users ($M = 128.85s, SD = 66.03s, N = 20$).

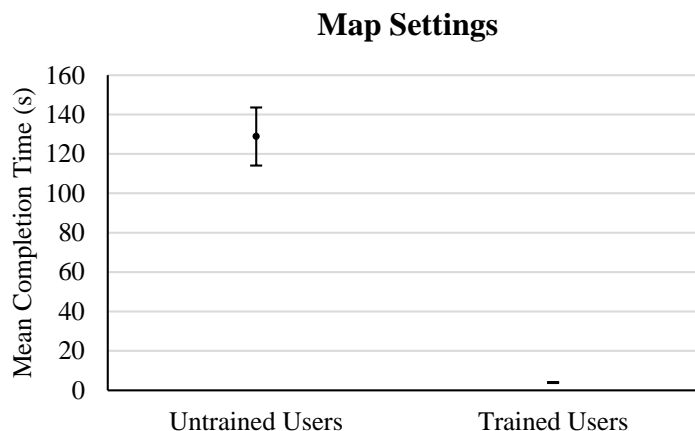


Figure 8. Plots of task completion time for the map settings

3.1.4. Survey Comparison of Weather Alert Messages

3.1.4.1. Comparison between the Original and Proposed STW Messages by Both Untrained and Trained Users

A paired sample *t* test was conducted to evaluate whether a statistical difference existed between the mean rating scores of each pair of statements in the original and proposed STW messages when evaluated by untrained and trained users. The results revealed that there was a very significant mean rating scores difference in each pair of statements for both groups of users. Specifically, Table 5 shows the paired *t*-test results in terms of mean (*M*), standard deviation (*SD*), standard error (*SE*), test statistics (*t*), degree of freedom (*DF*), and *p*-value (*P*).

Table 5. Paired *t*-test results for the comparison between NWS STW message and proposed STW messages by untrained users and trained users

	Untrained Users						Trained Users					
	<i>M</i>	<i>SD</i>	<i>SE</i>	<i>t</i>	<i>DF</i>	<i>P</i>	<i>M</i>	<i>SD</i>	<i>SE</i>	<i>t</i>	<i>DF</i>	<i>P</i>
NWS Header	2.00	.86	.19	18.29	19	<.001	2.75	1.25	.29	21.15	19	<.001
Proposed Header	8.75	1.07	.24				9.15	.81	.18			
NWS Delimiter	3.30	2.18	.49	11.08	19	<.001	4.10	1.02	.23	17.35	19	<.001
Proposed Delimiter	8.65	1.09	.24				9.05	.83	.18			
NWS Letter Format	2.90	1.25	.28	10.89	19	<.001	3.75	1.07	.24	17.25	19	<.001
Proposed Letter Format	8.45	1.32	.29				9.10	.79	.18			
NWS Message Satisfaction	2.75	1.02	.23	15.29	19	<.001	3.35	.67	.15	29.52	19	<.001
Proposed Message Satisfaction	8.50	1.10	.25				8.80	.62	.14			

3.1.4.2. Comparison between the Untrained and Trained Users for Both Original and Proposed STW Messages

An independent sample *t* test was conducted to determine whether a statistical difference existed between untrained and trained users on each comparison item in both the NWS STW message and the proposed STW message. The results showed that both groups of users rated all the messages comparison items similarly (*p*-values >.05), except for NWS Header, NWS Letter Format, and NWS Message Satisfaction (*p*-value <.05) [see Table 6].

Table 6. Independent t-test results for the comparison between untrained and trained users for both original and proposed Severe Thunderstorm Watch (STW) messages

	Untrained Users vs. Trained Users	
	Test Statistics (<i>t</i>)	<i>p</i> -value
NWS Header	2.21	.03
Proposed Header	1.33	.19
NWS Delimiter	1.49	.15
Proposed Delimiter	1.31	.20
NWS Letter Format	2.31	.03
Proposed Letter Format	1.90	.07
NWS Message Satisfaction	2.20	.03
Proposed Message Satisfaction	1.06	.29

3.1.4.3. Comparison between the Original and Proposed WA Messages by Both Untrained and Trained Users

A paired sample t test was performed to determine whether a statistical difference existed between the mean rating scores of each pair of statements in the original and proposed WA messages when evaluated untrained and trained users. It was revealed that there was a very significant mean rating scores difference in each pair of statements for both groups of users. Specifically, Table 7 shows the paired t-test results in terms of mean (M), standard deviation (SD), standard error (SE), test statistics (t), degree of freedom (DF), and p-value (P).

Table 7. Paired t-test results for the comparison between NWS STW message and proposed WA messages by untrained users and trained users

	Untrained Users						Trained Users					
	<i>M</i>	<i>SD</i>	<i>SE</i>	<i>t</i>	<i>DF</i>	<i>P</i>	<i>M</i>	<i>SD</i>	<i>SE</i>	<i>t</i>	<i>DF</i>	<i>P</i>
NWS Information Location	3.05	1.39	.31	12.33	19	<.001	3.00	1.30	.29	17.25	19	<.001
Proposed Information Location	8.35	1.07	.23				9.00	.86	.19			
NWS Word Expressions	2.65	1.09	.24	11.50	19	<.001	2.75	.97	.22	30.44	19	<.001
Proposed Word Expressions	8.15	1.50	.33				9.15	.75	.17			
NWS Terminology	2.55	1.00	.22	15.16	19	<.001	2.70	1.29	.25	20.68	19	<.001
Proposed Terminology	8.70	1.34	.30				9.15	.75	.17			
NWS Message Satisfaction	2.70	.66	.15	21.09	19	<.001	3.05	1.15	.26	24.72	19	<.001
Proposed Message Satisfaction	8.80	1.01	.22				8.75	.55	.12			

3.1.4.4. Comparison between the Untrained and Trained Users for Both Original and Proposed Wind Advisory (WA) Messages

An independent sample t test was conducted to examine whether a statistical difference existed between untrained and trained users on each comparison item in both the NWS WA message and the proposed WA message. The results illustrated that both groups of users rated all the messages comparison items similarly (p-values >.05), except for proposed information location and proposed word expressions (p-value <.05) [see Table 8].

Table 8. Independent t-test results for the comparison between untrained and trained users for both original and proposed Wind Advisory (WA) messages

	Untrained Users vs. Trained Users	
	Test Statistics (<i>t</i>)	<i>p</i> -value
NWS Information Location	.35	.73
Proposed Information Location	2.10	.04
NWS Word Expressions	.31	.76
Proposed Word Expressions	2.68	.01
NWS Terminology	.45	.66
Proposed Terminology	1.31	.20
NWS Message Satisfaction	1.19	.24
Proposed Message Satisfaction	.20	.85

3.2. Exit Survey Results

3.2.1. 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).

3.2.2. The Overall Usability of the Weather Radio Application

The results showed that the majority of participants rated the overall usability of the Weather Radio application 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 application as “Good”. In addition, 25 % of the participants (10 participants) rated the usability of the application as “Poor”, while no extreme ratings were reported.

3.2.3. Comments on the Usability of the Weather Radio Application

Table 9. Content analysis of usability comments from both untrained and trained users

Category	Problem/Expectation	No. of Comments	Representative Examples
Settings	Map settings	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	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	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)”

The responses to this question were analyzed using the qualitative content analysis technique [See Table 9]. 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 considered 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, which they did not know the meaning of “NWS”.

The analysis also revealed that twenty comments were included as issues in location search task. Fourteen of the comments were about the difficulty of locating a specific place on the map as the map was full of places and the display was relatively poor. Six comments were 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 application. 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.

4. Discussion and conclusion

The findings from this experiment revealed multiple usability issues that were associated with the tested features of the Weather Radio application. Those issues could also be found in several other weather alert applications and in applications with similar inherent features. The issues, implications, and the proposed solutions are discussed below.

For the location search task, both untrained and trained 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 2.4.1 for details]. In addition, a counter-intuitive 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 mobile phone displays. This implies that using the pin feature may not be the most efficient option when searching for a location on a mobile 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. Therefore, the pin feature may be more useful in this situation, especially if guidance based on contextual information is provided. Thus, it is recommended to include both features in weather alert applications, as well as in other mobile applications with embedded location search features.

The alert settings task was problematic to untrained users compared to trained 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 4 (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 4 (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 jargons and unclear abbreviations to enhance the user’s experience. These recommendations are also believed to be useful for non-weather applications as understandability of displayed information (Panach et al., 2008) and inclusion of the least amount of information required for accessing the features (Whitenton, 2016) are among the top usability requirements.

The map settings task was extremely challenging for untrained users. This was obvious as the untrained users needed substantially longer time to complete this task compared to trained users. In addition, the qualitative content analysis showed that most of the users’ concerns and comments regarding the usability of the application’s features were on the map settings task. The issues with this task were attributed to the included counter-intuitive 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 map itself, within the home screen, was tapped. Based on the untrained users’ responses to the exit survey and 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 the entire application. Hence, it is recommended to include the map settings menu in the home screen menu instead of its current hidden location; perhaps in the general settings menu when the home screen Gear icon is tapped [see Figure 4 (a) to see the Gear icon]. In addition, the application’s developers should consider creating a better representative icon of the map settings menu and keeping the “i” icon for displaying information about the application. 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 trained and untrained 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. Hence, it would be beneficial if NWS considered sending user friendly alert messages or providing a guideline that allows weather application 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 mobile weather system, in a timely manner may significantly impact users’ lives.

5. Limitation and Future Research

In this experiment, two important attributes were used: efficiency of use (objective term) and user satisfaction (subjective term). As explained in the introduction, there could be other important attributes to consider. However, those other attributes are not necessarily orthogonal among one another. For example, learnability can be highly correlated with efficiency and satisfaction. In future research, we will further investigate the usage of multiple attributes with a particular focus on examining the level of interdependency among these attributes.

Another research idea is to consider modern usability approaches, such as the eye tracking approach, which accounts for the users’ cognitive and decision-making processes (Kang & Landry, 2015). In particular, the eye tracking tool, in addition to the traditional usability objective attributes such as efficiency, adds more quantitative evidences and increases content validity. For instance, using the eye tracking approach, Dros et al. (2015) examined how people interact with weather forecasts shown on television.

Finally, a future study regarding the weather alert systems could test the effect of different screen sizes such as iPads, tablets, and desktop computer displays based on the performance of users. The findings from this study may determine the importance of the screen size factor for usability evaluation.

References

- Ahmadi, H., & Kong, J. (2008). Efficient web browsing on small screens. In *Proceedings of the working conference on Advanced visual interfaces* (pp. 23-30). ACM.
- Alluri, A. (2013). 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
- Baharuddin, R., Singh, D., & Razali, R. (2013). Usability dimensions for mobile applications- A review. *Research Journal of Applied Sciences, Engineering and Technology*, 5, 2225-2231.
- Drogalis, T. J., Keyes, E., & Dhyani, L. (2015). The Weather Channel application usability test. Retrieved from <http://elkeyes.com/wp-content/uploads/2015/05/WeatherChannelUTResults.pdf>.
- Drost, R., Trobec, J., Steffke, C., & Libarkin, J. (2015). Eye tracking: evaluating the impact of gesturing during televised weather forecasts. *Bulletin of the American Meteorological Society*, 96(3), 387-392.
- Frekjm, E., Hertzum, M. & Hornbaek, K. (2000). Measuring Usability: Are Effectiveness, Efficiency, and Satisfaction Really Correlated? Retrieved from http://www.diku.dk/~kash/papers/CHI2000_froekjaer.pdf
- 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). ACM.
- 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

- International Standards Organization, (1997). *ISO 9241: Ergonomic requirements for office work with visual display terminals (VDTs)-Part 11: Guidance on usability*. Geneva, Switzerland: International Standards Organization.
- 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.
- Kang, Z. and Landry, S.J. (2015). An eye movement analysis algorithm for a multi-element target tracking task: Maximum transition-based agglomerative hierarchical clustering. *IEEE Transactions on Human-Machine Systems*, 45(1), 13-24.
- Khanum, M. A., & Trivedi, M. C. (2012). Take care: a study on usability evaluation methods for children, *International Journal of Advanced Research in Computer Science*, 3(2), 101-105.
- Krueger, R. A., & Casey, M. A. (2002). Designing and conducting focus group interviews. *Social Analysis, Selected Tools and Techniques*, 4-23.
- Nayebi, F., Desharnais, J. M., & Abran, A. (2012). The state of the art of mobile application usability evaluation. In *Proceeding of the 25th Annual IEEE Canadian Conference on Electrical and Computer Engineering* (pp. 1-4), IEEE Canada.
- Nielsen, J. (1994). *Usability engineering*. Mountain View, CA: Elsevier.
- Nielsen, J. (2012). How Many Test Users in a Usability Study? Retrieved from <https://www.nngroup.com/articles/how-many-test-users/>.
- Panach, J. I., 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 Second ACM-IEEE international symposium on Empirical software engineering and measurement* (pp. 354-356). ACM.
- Poole, A., & Ball, L. J. (2006). Eye tracking in HCI and usability research. *Encyclopedia of human computer interaction*, 1, 211-219.
- Preece, J., Rogers, Y., Sharp, H., Benyon, D., Holland, S., & Carey, T. (1994). *Human-computer interaction*. Essex, UK: Addison-Wesley Longman Limited.
- Saleh, A. M., & Ismail, R. B. (2015). Usability evaluation frameworks of mobile application: a mini-systematic literature review. In *Proceeding of the 3rd Global Summit on Education GSE* (pp. 231-239), Global Summit on Education GSE.
- 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* (p. 70). ACM.
- Shackel, B. (1991). Usability-context, framework, definition, design and evaluation. *Human factors for informatics usability*, 21-37.
- 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). ACM.
- 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.
- Statistica (2017). Number of smartphone users in the United States from 2010 to 2020. Retrieved from <http://www.statista.com/statistics/201182/forecast-of-smartphone-users-in-the-us/>.
- van Elzakker, C. P., Delikostidis, I., & van Oosterom, P. J. (2008). Field-based usability evaluation methodology for mobile geo-applications. *The Cartographic Journal*, 45(2), 139-149.
- Weather Decision Technology (2016). Retrieved from <http://wdtinc.com/>.
- Weather Radio (2016). Retrieved from <http://weatherradioapp.com/>.
- Whitenton, W (2016). Website Forms Usability: Top 10 Recommendations. Retrieved from <https://www.nngroup.com/articles/web-form-design/>.
- 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.