User-Centered Design (UCD) of Time-Critical Weather Alert Application

Abdulelah M. Ali[1], Abdulrahman Khamaj[2], Ziho Kang[3], Majed Moosa[4], Mohd Mukhtar Alam[5]
Department of Industrial Engineering, Jazan University, Jazan, Saudi Arabia[1, 2, 4]
School of Industrial and Systems Engineering-College of Engineering, University of Oklahoma, Oklahoma[4]
Department of Mechanical Engineering, Vivekananda Global University, Jaipur, Rajasthan, India[5]

***Abstract*—Weather alert applications can save precious lives in time-critical risk situations; however, even the most widely used applications may fall short in intuitive interface and content design, possibly due to limitations in the users participation in the design process and in the users range considered. The objective of this study was to investigate whether the application of UCD principles and usability guidelines can improve the use of and satisfaction of time-critical weather alert apps by the public and or expert users. A prototype of a UCD-based weather alert application was developed and evaluated. Initially, thirty-two voluntaries participated in the identification of the important features that lead to the development of the porotype, and then the prototype was tested with another eighty participants (40 young and 40 elderly). The prototype includes five enhancements: auto-suggested location search, an all-inclusive interface for weather forecasts, message alert, visual and intuitive map settings, and minimalism-oriented alert settings. The enhanced functionality was compared to similar functionality in existing commercial weather applications. Effectiveness (completion rate, error count, error severity, and error cause), efficiency (time to completion), and satisfaction (post-task and post-test surveys) were measured. The results showed the enhancements significantly improved performance and satisfaction across both age groups compared to equivalent functionality in the existing app. The Mann-Whitney U test showed a statistically significant difference (p<0.001) in task satisfaction and number of errors between the two apps for all tasks. The Mann-Whitney U test showed a significant difference (p<0.001) in the across all tasks between the two apps Also, overall, young people with existing apps outperformed elderly, and both young and elderly with enhanced apps performed very high. Therefore, the enhancements implemented through the UCD process and usability guidelines significantly improved performance and satisfaction across both age groups to facilitate timely action necessary during a crisis.**

Keywords*—User-centered design; time-critical weather alert apps; weather forecasts; map set-tings; message alert

I. Introduction

With the advancement of smartphone technology, countless people have started to receive weather information on the smartphone platform. Due to the requirement for immediate action weather apps are important and potentially life-saving tools in time-critical severe weather events such as flooding and tornadoes. Specifically, usability issues with a poorly designed weather app interface can disproportionately affect certain end-user groups, especially the elderly. Elder users often experience problems with age, such as decreased working memory, and decreased visual, motor abilities, difficulties and cognitive [1-3]. In addition to physical and cognitive difficulties, elderly often struggle to cope with the rapid development of smartphone technology [4]; however, smartphone usage among older users continues to grow. As of 2017, nearly 74% of the U.S. population aged 50 to 64 and 42% of the population aged 65 and older owned a smartphone [5]. Numerous studies have focused on age differences in the use of smartphone apps in various fields, including information technology [6], healthcare [7], and communication [8]. However, to our knowledge, there are limited studies on usability testing of weather apps [9, 10], and no one has applied user-centered design (UCD) to include a wide range of users, especially elderly.

Drogalis et al. [10] identified usability issues with weather alert applications, such as the inclusion of hidden map menus that required prior counterintuitive actions and the lack of feedback on actions that were performed. Khamaj and Kang [9] investigate usability issues such as poor visualization of critical weather information, inappropriate language usage in time-critical alert messages, and inefficient location search functionality. However, neither of the two studies mentioned above addresses the question of how to effectively meet users' issues and needs. One possible way to address usability issues is to apply a user-centered design (UCD) approach, especially when it comes to end-users of different age groups. In detail, UCD refers to a cycle of design stages in which developers consider the needs, capabilities, and constraints of target users at each stage [11]. Norman [12] and Mao et al. [13] established a general UCD framework and applied to the human factors by Schnall et al. [14]. Careful consideration of each stage is believed to result in an interface that is easy to use and deemed useful [15]. The purpose of this study is to investigate whether the UCD approach could address usability issues for young and elderly when interacting with a smartphone weather alert application.

II. Method

The specific contribution of this research is to (1) characterize and classify the issues and needs of users with a broad age range in the design of weather alert applications, and (2) relate these issues and needs to usability guidelines (or principles), (3) design a prototype based on these guidelines, and (4) evaluate the prototype by selecting a target group (i.e., young and elderly). The results could impact the entire design process of a weather alert app (or other similar apps) that could save lives in a time-critical event.

Fig. 1 provides a summary of the framework used and applied by the afore-mentioned researchers. Following the framework provided in Fig. 1, the first stage of the UCD process is to describe the user's issues and needs when interacting with an existing application. Issues and requirements can be aligned with major usability guidelines, such as those presented by Nielsen [16] and Rogers et al. [17] and smartphone app design specifications, such as those presented by Gove [18]. At this stage, a broad range of users can be recruited to investigate various problems and needs that they can characterize and classified.

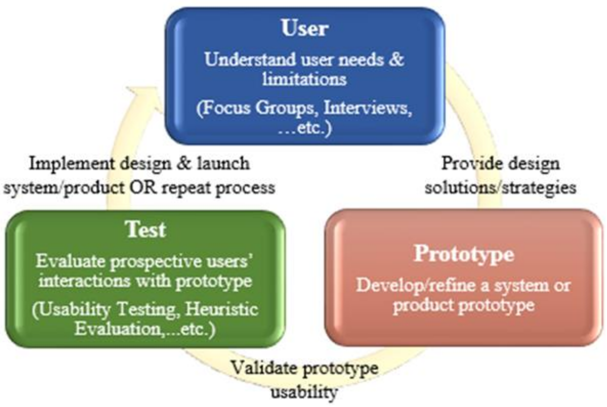


Fig. 1. Summary of the user-centered design framework (Jung et al. [15])

Specific to the weather app, questions and needs may relate to any of the features available on any typical weather app, such as location search, weather forecast, alert messages/notifications, alert settings, and radar/maps. The analyzed data can be characterized and classified using content analysis methods. One possible approach is thematic analysis [19]. This approach has been applied to critically evaluate and describe datasets, as well as to implicitly and explicitly identify provided ideas in form of codes refined as themes and sub-themes [20, 21].

The next stage is to develop a prototype based on the issues and needs of the target end users. For example, if a user asks about the inability to obtain weather forecasts effectively, especially in severe weather, and the need to easily map such forecasts with the locations they add; then these can be design and is implemented by adding location associations to the home screens, which can be accessed with a swipe whenever users open their weather app.

The final stage is to evaluate the developed prototype. In particular, we can con-sider the effectiveness of the prototype by targeting specific populations such as the older. In this case, we need to consider at least two factors: the type of application (i.e. existing vs. new prototypes) and the age group (i.e. young vs. older). Additionally, usability testing can be conducted using several qualitative and quantitative measures, such as completion rates, error counts, error severity ratings, error reasons, completion times, and satisfaction surveys.

III. UCD Phase 1: Understand User Needs and Limitations

Participants: In this study, 32 participants were recruited voluntarily, were (32 ± 6.4) years (mean ± SD), and ranged in age from 18 to 47 years. Participants were students at the University of Oklahoma and regular users of 11 popular smartphone weather apps running on different operating systems.

Design: The participants were guided by skilled moderators and assistant moderators. Several topics related to user perceptions and behaviors of smartphone weather application usability are discussed. These questions are intended to capture general information: (for example, trends in downloading specific weather apps, prioritization of weather apps in critical and non-critical time weather conditions, and positive and negative usability understandings of weather apps), and specific information: (for example, using different location search methods, controlling alert settings, content and display of critical alert messages, and using menu icons and labels).

Apparatus: A Nikon D3200 camera was used to record the sessions with the participants. A desktop computer, a projector device, and a large whiteboard surface were used to display the questions to the participants.

Procedure: Participants were held in a well maintained controlled environment at the University of Oklahoma. Participants signed a consent form upon arrival and then described the purpose of the study. After that, discussions begin, with each session lasting approximately 90 minutes.

Data analysis: Data was collected through video recording sessions. Afterwards, the collected data was transcribed, characterized and classified using qualitative thematic analysis [19]. Specifically, using thematic analysis methods, the data are: 1) transcribed verbatim, 2) encoded using representative words and phrases, and 3) re-fined into common themes.

A. *Results of Phase 1*

Participants presented three main themes: usage efficiency, user cognitive load, and effectiveness. These themes are primarily associated with common features on most popular weather apps (location search, weather forecast, alert messages, alert settings, and radar maps), as detailed below.

1) Usage efficiency: Participants were highly concerned with the time and number of steps required to access time-critical information, such as in a weather app. Participants shared several examples of inefficiencies in the design of the location search feature in popular weather apps. For example, in addition to conventional search methods (such as zip and/or city), the widely used weather app Weather Radio has recently adopted a precise location search feature (which enables users to search and save the specific locations) to provide users with accurate weather forecast. The application restricts users from searching for the different locations as required to navigate on the map and identify the desired location (see the example in Fig. 2(a)). Therefore, end users need an effective search method, since this feature needs a great degree of knowledge of the geographic area of the map, widespread visual attention, and repeated zooming in/out inside the small smartphone screen. Another efficiency requirement has to do with weather forecasting (such as humidity and temperature) feature. Users indicated that a limited number of steps were required because they felt it was accessed more frequently than other features. Several current weather applications involve multiple steps on several screens to access weather forecasts (see examples in Fig. 3(a-d)).

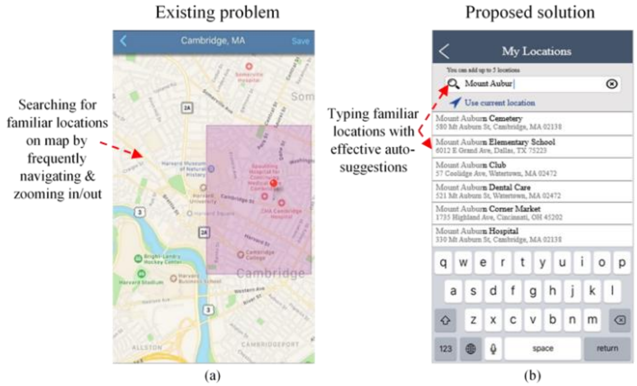


Fig. 2. Screenshot example of a user's main problem with the location search feature and proposed solution. (a) Screenshot of location search in the Weather Radio app, which has efficiency issues where users spend a lot of time navigating and zooming in/out on the map to find familiar locations. (b) Screenshot of a proposed efficiency solution in the UCD app, where the user simply types a familiar location and selects it from a list of auto-suggestions

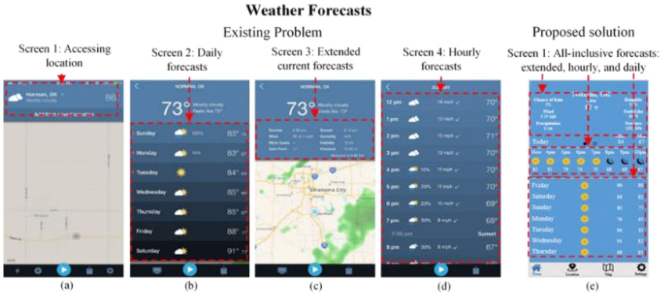


Fig. 3. Screenshot example of a user's main problem with the weather forecasts feature and proposed solution. Screens (a) to (d) are screenshots of weather forecasts in the Weather Radio app with efficiency issues: multi-screen/step forecasts, where the user must access each set of forecasts from a different screen. Screen (e) is a screenshot of the proposed efficiency solution in the UCD app, where users can access predictions for all locations within one screen immediately after opening the app

2) User cognitive load: Due to critical situations associated with extreme weather conditions, users also expressed a need for weather apps to simplify their cognitive processes and decision-making strategies. For example, users place high emphasis on the requirement for concise and structured push alert messages throughout severe weather conditions; alerts are created by the weather agency's systems and automatically directed to third parties, including weather apps, where they push them to the end user exactly as they are received. This is because alerts often contain technical data (such as geographic area codes) and cluttered data (see the example in Fig. 4(a)), which can hinder users' understanding of the alert and rise their cognitive load. Additionally, alerts often contain a huge quantity of evidence; most of this is not related to user-saved locations (e.g. names and information for all alert areas). While retrieving information about distant or isolated locations is critical for some operators in certain usage contexts, such as when traveling, making them a major part of alerts can be detrimental to users. Users also need a radar map feature that is built-in to identify and understand. Participants shared several examples of few popular weather applications that lack intuitive and/or visible instructions on in what way to access maps of the respective saved localities or control their settings (see examples in Fig. 5(a, b)).

3) Effectiveness: Owing to the limited smartphones screen size and the critical information displayed, users demand optimized and flexible functionality. For example, in the alert settings feature, the weather app either lets users control of all types of weather alerts and sub- alerts (see example in Fig. 6(a, b)), or does not permit the control of any alerts, but automatically uses active warnings as alerts. For this reason, users have expressed concern about large numbers of weather warnings, most of which are rarely needed and / or insignificant for the average user. Additionally, notifications that push any active alerts without end-user control are considered mandatory interactions. Therefore, users have expressed the need for the ability to control only a few relevant alerts.

IV. UCD Phase 2: Develop Product Prototype

A prototype weather app was developed after understanding the specific concerns and needs of users and taking into account widely used smartphone design heuristics, age-related constraints and needs, and common usability guidelines. The prototype application was designed using the InVision application software (https://www.invisionapp.com/company) called "EZ Weather".

A. Design Proposal Overview

The proposed design solution for the weather app features has been implemented in the EZ Weather app. To address the location search problem, a similar approach to the Google Maps app could greatly improve the efficiency of this feature: enter a familiar location with effective auto-suggestions. With this feature, the user can add an exact location by typing the location name/address (e.g. hospital name/address) in the search bar of the application and selecting it from the auto-suggest list (see example in Fig 2(b)). To reduce the time and number of steps to access weather forecasts, it may be helpful to have all weather forecasts for each saved location in one screen: all-inclusive weather forecast in one screen (see example in Fig 3(e)). For alert messages, a method for filtering the content of the message can significantly reduce the cognitive load on the user by including only relevant and necessary information on the main alert screen that is relevant to the user's saved location; all other information can be accessed from the secondary menu, including distant under-alert area.

In addition, simple language and hierarchically structured information using everyday language can enhance user understanding of messages and responses to alert threats. See Fig. 4(b)), for an example, content that is structured, prioritized, and language simplified. In addition, to facilitate the user's identification and understanding of the radar map feature, it may be helpful to display the list of all saved locations; from which the user can visually select the desired location; then the map of the selected location and its settings are displayed: use the visible and intuitive Map menu (see example in Fig. 5 (c & d)).

To avoid an excessive number of alerts that the user needs to control and interact with, and to give the user the freedom to control the push alert notifications, it may be beneficial to include the most critical and common alerts (on/off) to control. In bad weather, users will receive alert notifications about the time-critical alerts they have turned on (e.g. tornado warnings). Alerts that are not time-critical, such as wind, are not spontaneously sent to the user as notifications; these alerts only appear after clicking the respective symbol on the affected location screen (Fig. 6 (c & d)).

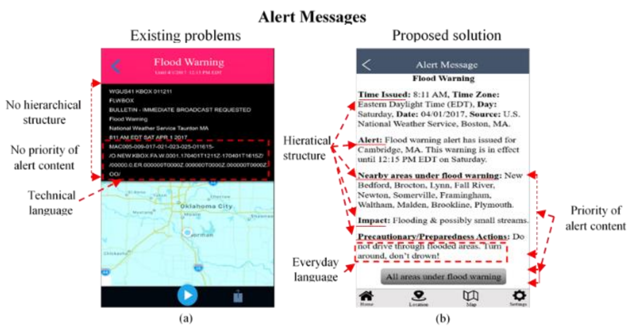


Fig. 4. Screenshot example of a user's main problem with an alert messages feature and proposed solution. Screen (a) is a screenshot of an alert message in the Weather Radio app, showing information about all alert-deficient areas within a scrollable area, uses code and technical language, and without hierarchical information. Screen (b) is a screenshot of the proposed solution in the UCD app, by prioritizing alert content (information about under-alert saved location and nearby areas on main screen and distant under-alert areas accessed from a secondary screen), using everyday language, and with structure hierarchical information

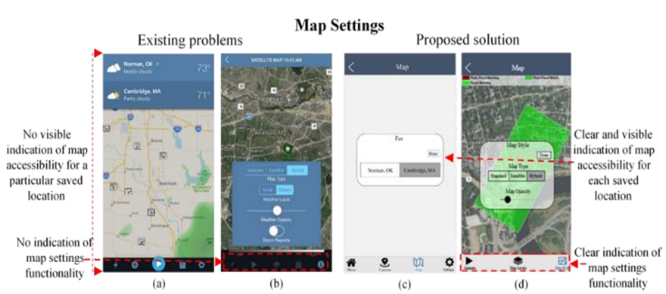


Fig. 5. Screenshot example of a user's main problem with radar/map feature and proposed solutions. Screens (a) & (b) are screenshots of radar/map in Weather Radio app that includes visibility and recognition issues, as neither indication is available on how to access map of a desired saved location as in (a) nor the functionality of icons is easily recognized as in (b). Screens (c) & (d) are screenshots of proposed solutions in the UCD app, where (c) shows all saved locations for a user to select and access map of a desired location and (d) provides icon labels to ease functionality recognition

Other key usability and smartphone application guidelines may also need to be considered to improve the overall user experience. These guidelines include labeling all icons with representative text labels, using short descriptive information next to menu options, providing appropriate feedback on user actions, and using appropriate colors with easily distinguishable contrast to visualize and indicate weather conditions (e.g. red screen for locations under a warning alert), provide shortcuts (for example, constant main icons across all screens), text in relatively large fonts, and use a consistent and intuitive layout.

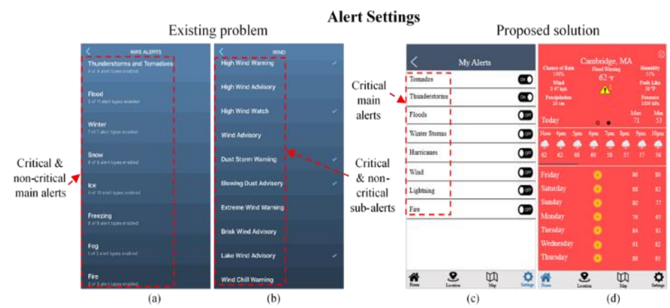


Fig. 6. Screenshot example of a user's main problem with alert settings feature and proposed solution. Screens (a) & (b) are screenshots of alert settings in the Weather Radio app, where users are required to scroll up/down and control the settings of all-weather types' main alerts and sub-alerts (e.g. wind alert has 16 sub-alerts), including critical and non-critical alerts. Screens (c) & (d) are screenshots of a proposed solution of minimalist alert settings in the UCD app, where users need to control the settings of only critical alerts as in (c); non-critical alerts' information can be directly accessed from screen of affected location (e.g. by tapping yellow alert icon as in (d))

V. UCD Phase 3: Evaluate Prospective Users' Interactions with Prototype

A. Method of Phase 3

Participants: In this study, eighty (40 young and 40 elder) regular users of the iOS smartphone weather app (first-time users of the tested app) were voluntarily recruited for the experiment. The younger and older participants were 25.9 ± 4.8 years (mean ± SD) and 57.4 ± 4.3 years, respectively, and ranged in age from 18 to 35 years and 50 to 66 years, respectively. The participants were assigned randomly to perform tasks in two test apps (Weather Radio and EZ Weather). Both applications were used by 40 users (20 young and 20 elder). Recruitment was based on personal communications, mass emails from universities and flyers hanging on doors of various public buildings. Participants signed a consent form upon arrival and then described the purpose of the study. The experimental protocol was approved by the ethical committee of the department of Human Research Participant Protection (HRPP), University of Oklahoma, Norman campus with IRB: 6681/2021.

Apparatus: For the present study, both the applications i.e., Weather Radio (version 3.0.5) and EZ Weather (version 1.0.0) are installed on iOS Smartphone. Using high fidelity simulations (InVision), a powerful interaction design system, we displayed recorded alert messages on the Weather radio at any time during the experiment. To record user interactions, in particular to calculate the time to complete a given task, to count and classify errors the Nikon L340 camera was used. Demographics, post-task and post-test surveys were printed on paper.

Scenario & Tasks: Participants were given a scenario in which their grandmother was admitted to Mount Auburn Hospital in Cambridge, MA, and noticed that a flood warning had been announced in the Cambridge area. To understand grandmother's exact location alert risk level and access all relevant information, participants can search for and add to grandmother's exact location, access and identify relevant weather forecasts and alert messages. Further, participants need to view alerts on map in map settings and adjust alert settings to receive relevant alert notifications. Fig. 7 shows the successful completion of the scenarios and tasks.

Procedure: First, participants were explained the purpose of the survey, signed a written consent form, and then completed the demographic survey. None of the participants received training and were not given the opportunity to practice themselves using the test app; the participants were given scenario and task guidelines and were asked to start the experiment after they were told they were ready. Weather Radio is a running application, so participants won't see any alert messages unless during an active alert, so Weather Radio's alert message assignment was presented to participants through a smartphone interface designed by InVision. The alert message displayed to participants is a flood warning posted to Weather app users in Cambridge, MA on April 1, 2017. After finishing each task, participants were asked to complete a post-task survey to provide their instant feelings about each task/ feature. Finally, participants were asked to complete a post-test survey to evaluate their satisfaction with the test app.

Experimental Design & Variables: Full factorial design (2 levels of app * 2 levels of age group) were used in the present study as independent variables. The dependent variables were effectiveness, efficiency, and user satisfaction. The control variable was user experience; all users in both age groups must have at least six months of experience with iOS smartphones and the Weather app (excluding the app tested). Error metrics are the reasons for errors made (usability issues), how often they occur, the proportion of users who make mistakes, and the severity level of the cause of the error. The error severity ratings used in this study are based on a rating scale proposed by Nielsen [23] (Table I).

Finally, Pearson correlation test (r) was also performed to define the relationship between the usability measures used in this study.

The post-task survey was done by using a Single Ease Question (SEQ) with 7-point Likert rating scale [22]. This question was used because it was found to be as effective as other complex measures of task difficulty, namely the Usability Magnitude Estimate (UME) and Subjective Mental Effort (SMEQ) questionnaires [23]. The post-test survey was done by using User Interface Satisfaction (QUIS) questionnaire [24].

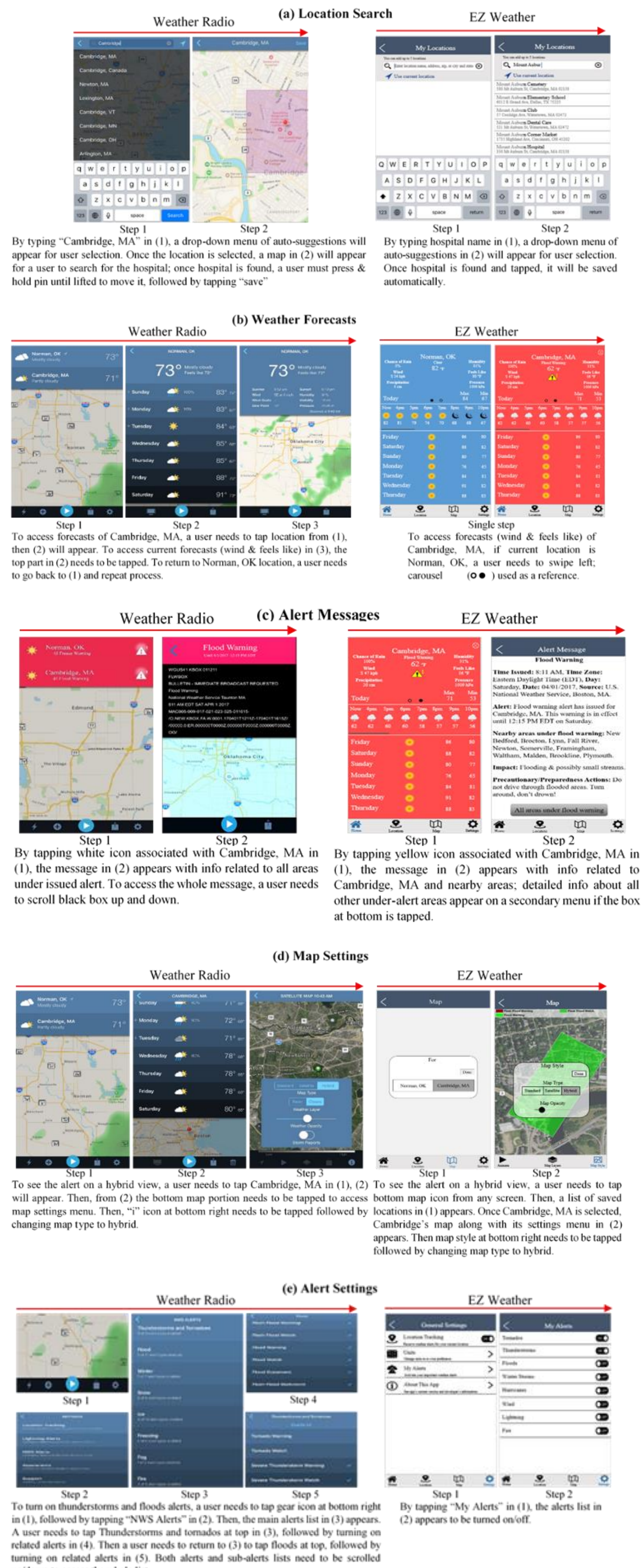


Fig. 7. (a-e). Process of performing tasks on Weather Radio vs. EZ Weather

Data Analysis: All data's were analyzed using SPSS version 23. A two-way analysis of variance (ANOVA) between subjects was performed to determine the effect of both application used and age group task completion time for all tasks. The Mann-Whitney U test was used to examine the size of the differences between the app and age group variables used, both in terms of the number of errors and the metrics of the post-task satisfaction survey. For post-test satisfaction surveys, mean and standard deviation errors were calculated.

TABLE I. NIELSEN'S SEVERITY RATING SCALE OF THE USABILITY PROBLEMS

Rating	Nature of 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

B. Results of Phase 3

1) Effectiveness

a) Completion rate: Fig. 8 shows that all users in both groups were able to complete a given task in EZ Weather successfully. However, several users were unable to complete all three tasks on Weather Radio, and older users had a higher failure rate on the location search task.

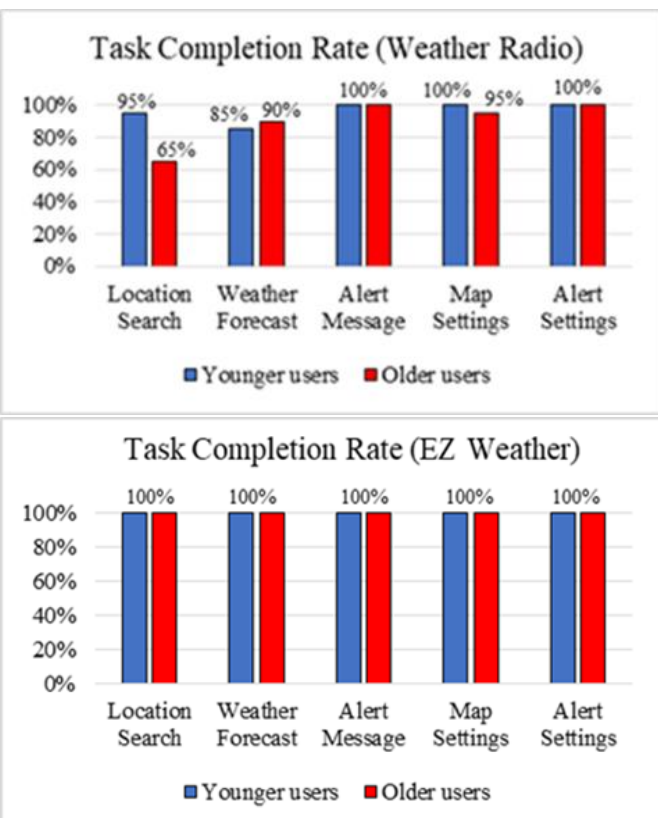


Fig. 8. Proportions of successful task completion rate for both app

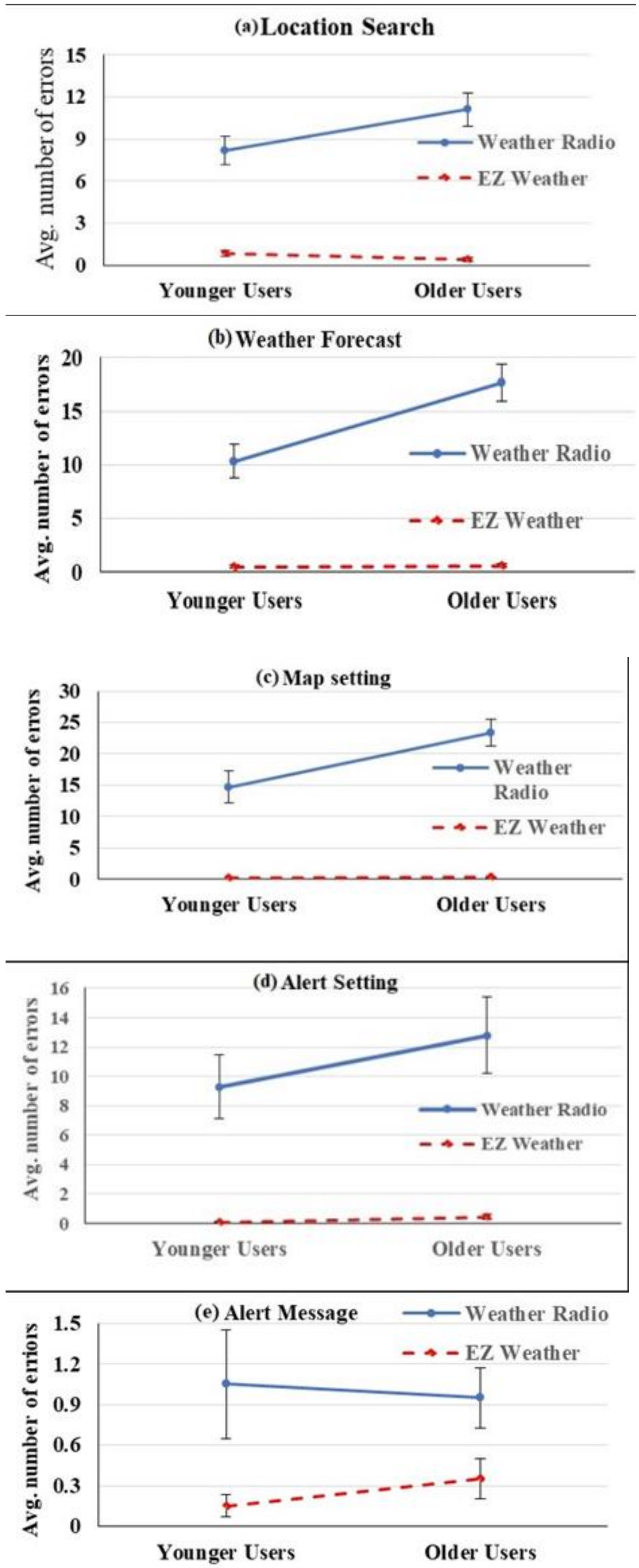


Fig. 9. (a-e). Average number of errors using both apps

TABLE II. MAN-WHITNEY TEST SUMMARY FOR NUMBER OF ERRORS

Task	Source	Z-score	U-test	P-value
Location Search	Age group	-1.22	677.50	0.223
	App used	-7.86	9	<0.001
Weather Forecast	Age group	-2.14	582	0.033
	App used	-6.64	123	<0.001
Alert Message	Age group	-.56	743	0.575
	App used	-7.68	20	<0.001
Map Settings	Age group	-2.42	655.5	0.015
	App used	-7.51	34.5	<0.001
Alert Settings	Age group	-.91	710.5	0.364
	App used	-8.01	10	<0.001

TABLE III. 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 issues		Average Severity Rating
		Young Users	Older Users	
Location Search	Users are having trouble finding locations and moving pins on the map.	178	253	3.5
Weather Forecast	Users couldn't easily find the weather forecast. The area that causes the prediction, if clicked, appears to be un-clickable.	231	366	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	17	4
Map Settings	Users struggle to start the task due to counterintuitive steps and an invisible map settings menu.	166	287	3
	Users are unaware of the functionality of the map settings icons because these icons are neither marked nor standardized across mobile applications.	126	194	2.5
Alert Settings	Users do not understand the function of home screen icons because these icons are neither marked nor standardized in smartphone applications. Also, the large number of alerts and sub-alerts seems to confuse users about the options required for the task.	186	256	3.5

b) Errors: Fig. 9(a-e) shows that errors on EZ Weather are significantly less than those on Weather Radio for all tasks, for both young and older users. The results showed that the older users had made more mistakes on both the apps compared to younger participants on all tasks except Weather Radio's alert message task and EZ Weather's location search task. The Mann-Whitney U test showed a significant difference ($p<0.001$) in the number of errors across all tasks between the two apps (Table II). The results showed no significant error differences for all tasks between the both groups, except for the weather forecast and map setting tasks. For other error-related metrics, almost all users in both groups made errors, and with different frequencies, due to the usability issues with every task in Weather Radio except the Alert Messages task (Table III). Nearly half of the participants made mistakes due to alert message usability issues. In contrast, fewer users made mistakes in EZ Weather's tasks, with a much lower frequency and average severity rating, compared to Weather Radio. Errors on EZ Weather are mainly caused by swipe actions (such as spelling mistakes and accidental clicks on adjacent function icons) (Table IV).

TABLE IV. 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		Average Severity Rating
		Young Users	Older Users	
Location Search	Users made typing errors when typing location name.	17	9	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	12	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	7	1
Map Settings	Users mistakenly tapped adjacent icons of unrelated functions.	4	6	0
Alert Settings	Users mistakenly tapped adjacent icons of unrelated functions.	2	9	0

2) Efficiency (task time): Fig. 10(a-e) shows that, on an average, the younger and older participants take fewer time to accomplish each task of EZ Weather compared to Weather Radio. The two age groups had relatively taken same times on all tasks for both apps. Therefore, two-way ANOVA (Table V) shows that on both weather apps, there were no significant time differences between young and elder users for all tasks except the map setting task. In the weather app tested, however, the time to complete all tasks was significantly different ($p<0.001$) between the both groups. In addition, there was no significant interaction between the both group and the app used for all tasks.

3) Post-task satisfaction ratings: Fig. 11(a-e) shows that both young and older participants found the EZ Weather task easier than the Weather Radio task. Younger users rated the ease of use for both apps higher compared to older participants, though the difference was not huge. In addition, the Mann-Whitney U test showed a statistically significant difference ($p<0.001$) in task satisfaction between the two apps for all tasks (Table VI). It also showed that young and older users were similarly satisfied with all tasks.

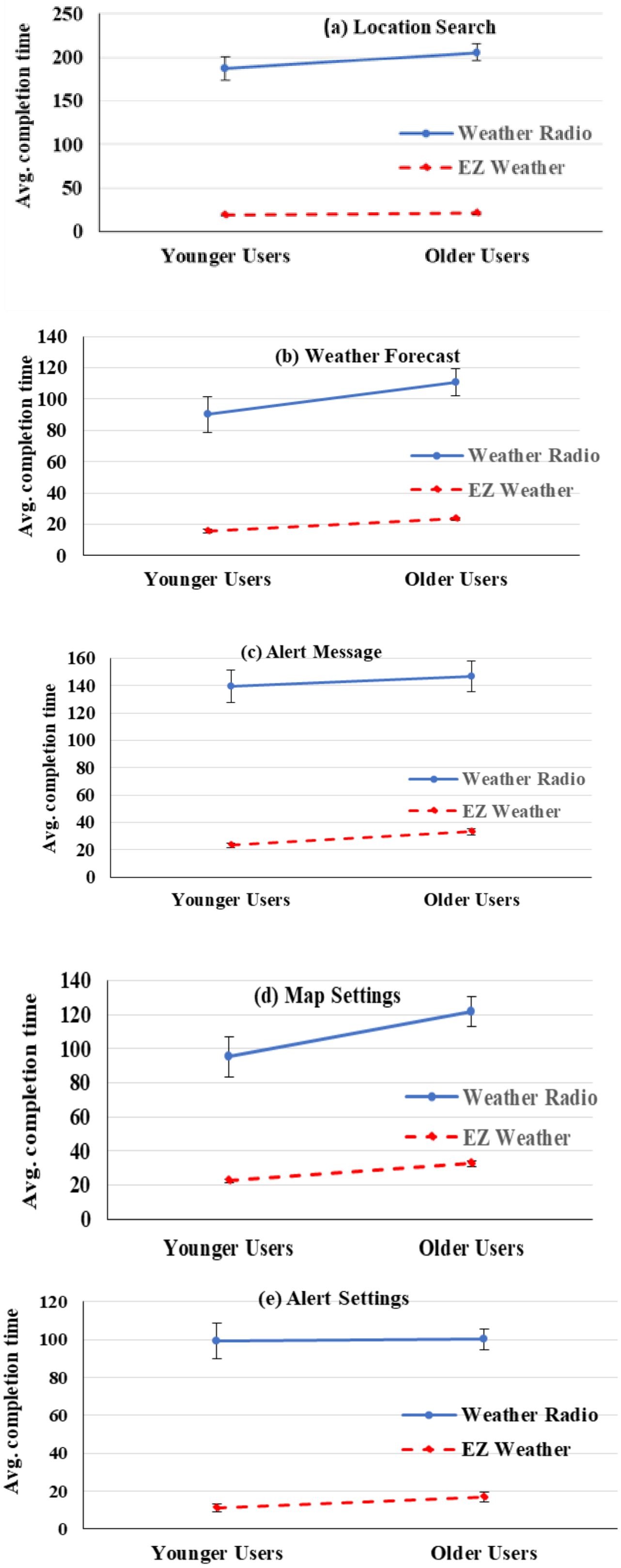


Fig. 10. (a-e). Average time spent (in seconds) in completion of tasks

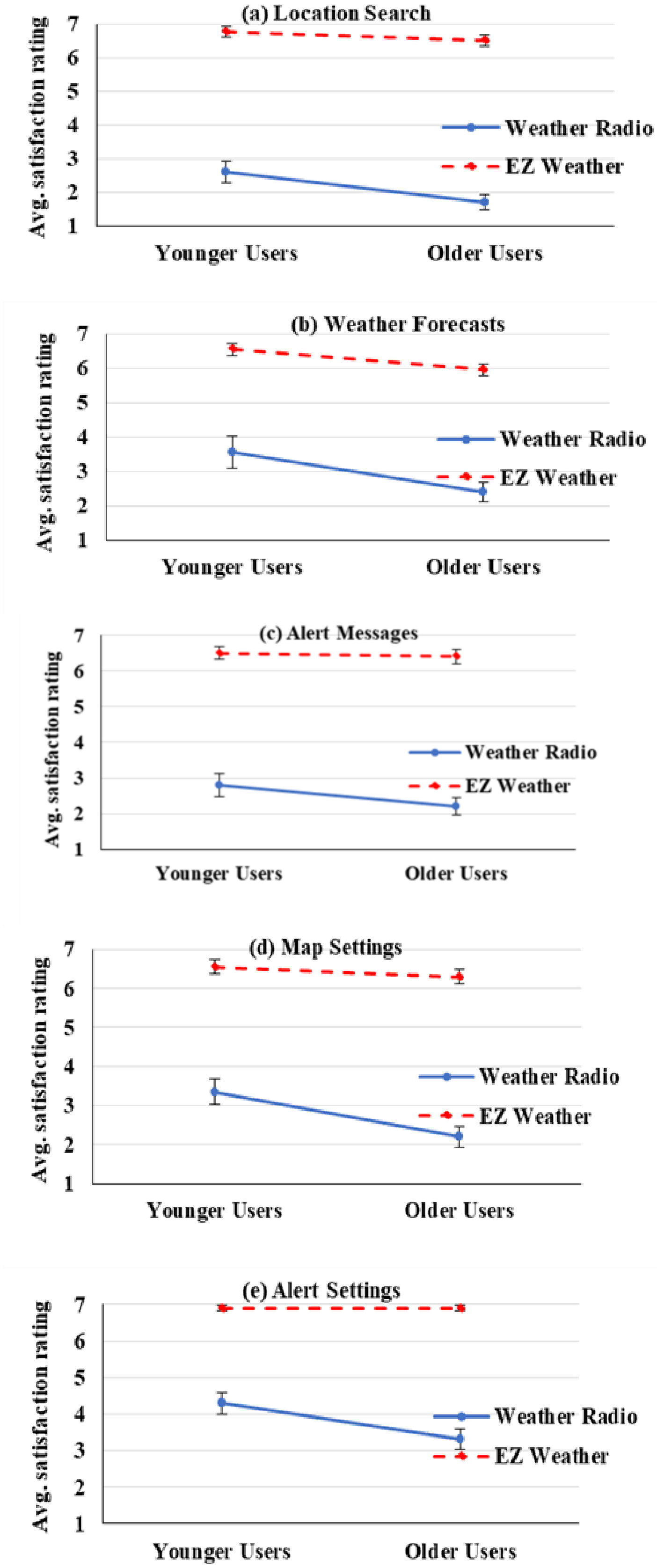


Fig. 11. (a-e). Average post-task satisfaction ratings of SEQ survey

TABLE V. TWO-WAY ANOVA SUMMARY FOR TASK TIME

Task	Source	F	P-value
Location Search	Age group	1.401	0.240
	App Used	445.554	<0.001
	User * App	0.949	0.333
Weather Forecasts	Age group	3.916	0.051
	App Used	127.384	<0.001
	User * App	0.805	0.373
Alert Messages	Age group	1.010	0.318
	App Used	189.608	<0.001
	User * App	0.023	0.879
Map Settings	Age group	6.012	0.017
	App Used	116.856	<0.001
	User * App	1.205	0.276
Alert Settings	Age group	0.332	0.566
	App Used	242.835	<0.001
	User * App	0.194	0.661

TABLE VI. MAN-WHITNEY TEST SUMMARY FOR POST-TASK SATISFACTION RATINGS

Task	Source	Z-score	U-test	P-value
Location Search	Age group	-0.298	769.5	0.766
	App used	-7.76	5	<0.001
Weather Forecast	Age group	-1.32	665	0.187
	App used	-7.58	25	<0.001
Alert Message	Age group	-0.842	727.5	0.401
	App used	-2.68	569	0.007
Map Settings	Age group	-1.33	667	0.184
	App used	-7.97	1.5	<0.001
Alert Settings	Age group	-1.32	668	0.185
	App used	-7.09	93.5	<0.001

4) Post-test satisfaction ratings: Overall, the results in Fig. 12 and 13 was based on the QUIS questions [24], where the same questions were asked to both groups of participants under same parameters, suggesting that younger and older participants feel comparatively similar to the interface de-sign specifications, and younger users are more satisfied. In addition, all participants were very pleased with EZ Weather, while most users of Weather Radio were disappointed. Users report that the screen design of EZ Weather is very good, the text and images are very clear, the fonts are very clear, the amount of information presented is sufficient, and the arrangement is reasonable. However, participants were appeared to be disappointed with the amount of information displayed on Weather Radio; they observed that there was too much information, particularly about time-critical alerts.

Users have found consistent terminology and messaging throughout EZ Weather as well as the location of on-screen instructions. Participants reported that performing any action produced predictable results in EZ Weather, with a very acceptable delay between actions. However, many participants noted that they were not sure what to expect when performing multiple actions in Weather Radio. Compared to Weather Radio, users report that using EZ Weather is very easy to pick up, learn, and efficient. Additionally, users seem to be very pleased with the number and order of steps required to accomplish each EZ Weather task. They also liked the feedback they received as they completed each task. Navigating features and remembering names and commands was also easy for participants in both groups. When it comes to the EZ Weather feature, although many users made no mistakes during the experiment, they reported that errors and typos were easy to correct. Users also like their ability to perform or undo actions using shortcuts. In addition, they say that users of any experience level can easily and consistently complete their tasks. The choices of color for both apps seem natural enough for most users, with EZ Weather having a higher user satisfaction rate.

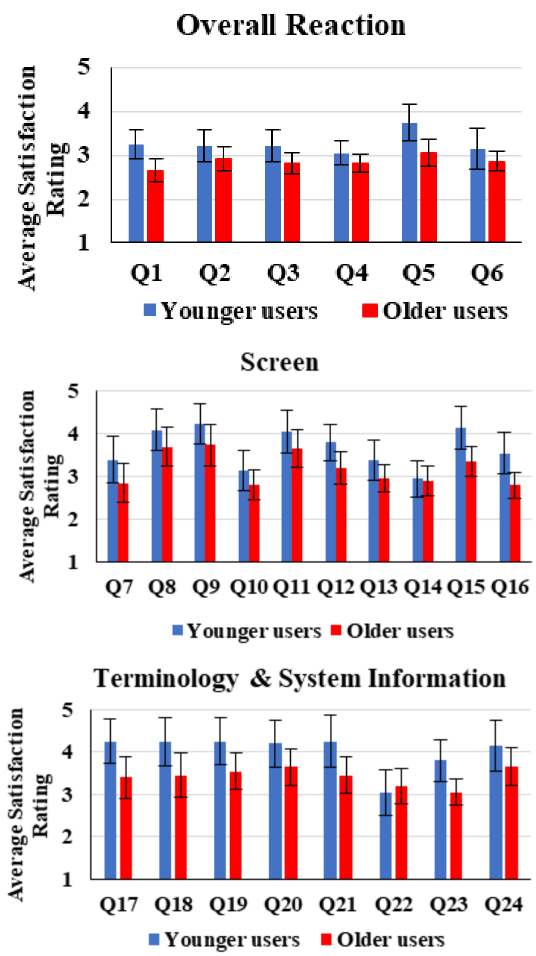
5) Correlations among usability measures: Tables VII and VIII show the Spearman rho correlation (ρ) results among all the indicators in this study. Since all users successfully performed a given task, the correlation between completion rates for the two age groups and other metrics on EZ Weather was not performed. The results showed a strong positive correlation between task time and errors, and between post-task and post-test satisfaction, for both age groups for both applications. However, the time spent on each task and errors were negatively correlated with post-task and post-test satisfaction scores.

TABLE VII. CORRELATION MATRIX FOR BOTH YOUNGER AND OLDER USERS ON WEATHER RADIO, IN TERMS OF P & (P-VALUES).

	Younger User			
	Time	**Errors**	**Task-Sat**	**Test-Sat**
Completion rate	-0.615 (0.002)	-0.682 (0.003)	0.766 (<0.001)	0.870 (<0.001)
Time		0.885 (<0.001)	-0.921 (<0.001)	-0.934 (<0.001)
Errors			-0.858 (<0.001)	-0.898 (<0.001)
Task-Sat				0.917 (<0.001)
Older User				
Completion rate	-0.651 (0.003)	-0.792 (<0.001)	0.784 (<0.001)	0.921 (<0.001)
Time		0.907 (<0.001)	-0.905 (<0.001)	-0.908 (<0.001)
Errors			-0.843 (<0.001)	-0.929 (<0.001)
Task-Sat				0.911 (<0.001)

TABLE VIII. CORRELATION MATRIX FOR YOUNGER AND OLDER USERS ON EZ WEATHER, IN TERMS OF P & (P-VALUES).

	Younger User		
	Errors	**Task-Sat**	**Test-Sat**
Time	0.798 (<0.001)	-0.883 (0.002)	-0.844 (<0.001)
Errors		-0.704 (<0.001)	-0.711 (<0.001)
Task-Sat			0.841 (<0.001)
Older User			
Time	**Errors**	**Task-Sat**	**Test-Sat**
Errors	0.781 (<0.001)	-0.820 (<0.001)	-0.851 (<0.001)
Task-Sat		-0.798 (<0.001)	-0.831 (<0.001)
			0.892 (<0.001)



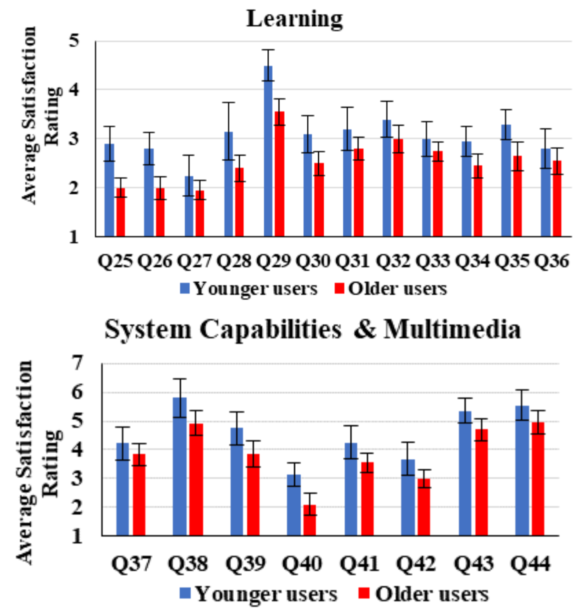
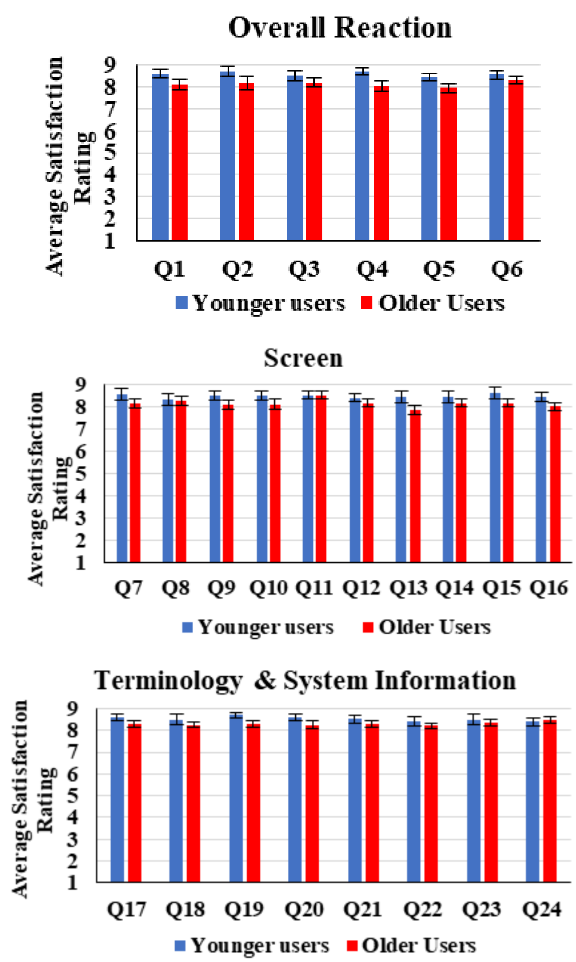


Fig. 12. Average post-test satisfaction rating of QUIS survey (Weather Radio)



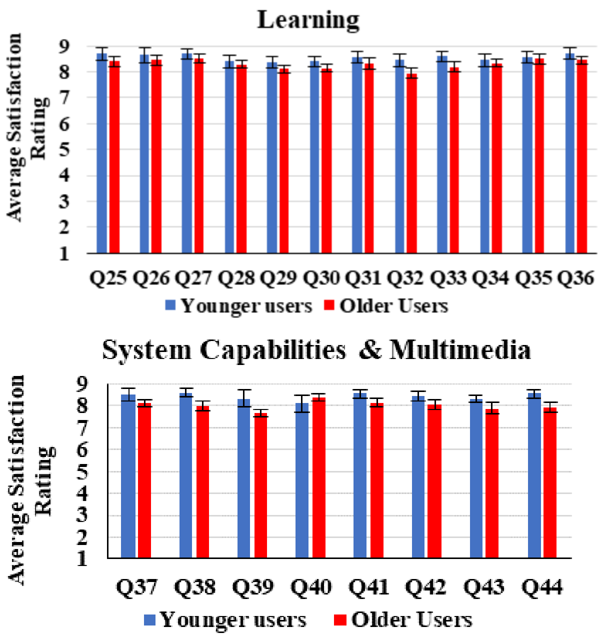


Fig. 13. Average post-test satisfaction rating of QUIS survey (EZ Weather)

VI. Discussion

Overall, this study establish that the UCD weather app (EZ Weather) was considerably more useful compared to the other popular weather app (Weather Radio) for all the evaluation metrics used. Also, younger and older participants seem to have comparable results on all tasks for both the weather application, with younger users experiencing marginally greater performance and satisfaction. These results show that regard-less of age, an application's interface design significantly affects end-user performance and perceptions (positive or negative) of application usability. The results further demonstrate that adopting a UCD approach to applications that contain time-critical data, such as weather applications, will result in extremely interactive and usable systems. Results for all metrics had showed that prioritizing and structuring critical evidence and using everyday terms throughout the EZ Weather interface greatly helps participants easily interacting with inherent features and finds them beneficial. For example, all young and older participants were able to complete alert message tasks in EZ Weather with far greater efficacy and task satisfaction compared to Weather Radio features.

Additionally, only a handful of users had errors in some of the usability issues category in EZ Weather's alert messages task, while nearly half of participants of each group had errors in this task for Weather Radio due to usability issues in the catastrophe category. In addition, the alert message task further demonstrate that the push alerts will greatly help end participants to correctly perceive alert threats and respond effectively to alert threats, particularly in the forthcoming weather conditions.

The greater usability of the EZ Weather is also due in part to consideration of efficiency principles during the design phase. This is supported by the very limited time and steps required to access any feature on EZ Weather. A good example is the weather forecast feature that includes all weather prediction information for respectively saved location within same screen (all-inclusive weather forecast); with just one swipe right or left, the user can access other saved locations' weather forecast. Although some users couldn't figure out the swipe functionality, from the first test, all success-fully completed the weather forecast task with significantly greater efficacy and task satisfaction, with fewer and less errors, than Weather Radio multiple-screen weather forecasts feature. Another example of an efficient feature of EZ Weather is the location search feature. With efficient auto-suggestions to type in familiar locations, both groups of users can complete the corresponding task in seconds. By contrast, navigating maps and pinpointing locations within the limited smartphone screen size is not only inefficient, but also ineffective, especially for older users, as demonstrated by Weather Radio's high failure rate and completion time for location searches.

Users' greater performance and satisfaction with EZ Weather than Weather Radio may also be related to EZ Weather's simple design. For example, limiting user-controlled alerts to the most common and critical alarms on EZ Weather (using a minimalist alarm setup) enables users to effectively and efficiently perform alarm set-up tasks with great satisfaction. Moreover, EZ Weather also permits participants to use non-critical alerts (via an illustrative icon displayed on the exaggerated location screen) during active alerts. The significant drop in performance and satisfaction with Weather Radio's alerts setting tasks was primarily because of the huge number of alerts and sub- alerts that users need to navigate and control (e.g., 16 sub- alerts for wind alerts alone). Two other important principles considered in the design of the EZ Weather interface are its usability and ease of recognition. For example, while the map setting task on both apps requires interaction with similar step and function icons. EZ Weather's greater performance and satisfaction on this task is supposed to be recognized to the visualization and intuitive menus (using the visualization and intuitive map menu), and appropriately labeled icons. These factors and results show that the actions have been performed with expectable results in a logical sequence. However, invisible elements in smartphone apps (such as the map settings menu bar) and icons that are neither marked nor standardized may be one of the main reasons for the dramatic drop in performance and satisfaction with Weather Radio tasks.

The usability and design principles monitored in the design of EZ Weather that may have ultimately contributed to great results include: feedback on system status (e.g. confirmation messages to perform activities), uniformity of application elements (e.g. settings menu), shortcuts usability to expedite interaction and correct errors, provide short descriptive information to help users understand the functionality of the corresponding feature, and use large text fonts along with high contrast and indicative colors to address age-related limitations for older users. The post-test satisfaction survey results were consistent with those of the task-based metrics. The two apps differed significantly in satisfaction with all interface standards across the two age groups, suggesting that applying usability guidelines to interface design not only results in high performance but also satisfaction for users. Also, on most

QUIS survey items, the two age groups are highly similar in satisfaction, meaning that age differences do not significantly affect user satisfaction. The strong correlations between all used metrics complement the results of previous study [25], which suggest that usability metrics are interdependent aspects that inform the usability of an interface. However, Frøkjær et al. [26] showed the dependence of usability metrics depends on whether the tested interface contains highly complex functionality. If the domain of interest includes complex features, a weak correlation is expected; if not, a strong correlation is most likely.

VII. Conclusion

This study validated the usability of the UCD weather app (EZ Weather) by comparing it to a popular app (Weather Radio). In addition, this study also considered young and older users to determine whether the usability of the test app would be highly influenced by age differences. Results for all metrics showed that users' performance and satisfaction on EZ Weather improved significantly regardless of age, demonstrating the importance of considering UCD methods, usability heuristics, and smartphone design rules in interface development.

Acknowledgments

The authors would like to thank all team members and support staff who were directly or indirectly involved during the development and testing of this application.

References

[1] P. Lawton. "Aging and performance of home tasks". Human factors. 1990, 32(5): 527-536.

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

[3] S. Czaja. "Factors predicting the use of technology: Findings from the center for research and education on aging and technology enhancement". Psychology and Aging. 2006, 21(2): 333–352.

[4] M. Khawaji. "Towards an enhanced user experience with smart phone weather alert applications: Usability and User-Centered Design approaches", 2018. Retrieved from https://hdl.handle.net/11244/302100.

[5] T. Levdikova. "Designing apps for elderly smartphone users", 2019. Retrieved from https://clutch.co/app-development/resources/designing-apps-for-elderly-smartphone-users

[6] F. García-Peñalvo, M. Conde and O. Matellán. "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, Heraklion, Greece, 2014.

[7] A. Morey, H. Barg-Walkow and A. Rogers. "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, Los Angeles, 2017.

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

[9] A. Khamaj, Z. Kang. "Usability evaluation of mobile weather hazard alert applications". Industrial and Systems Engineering Review. 2018, 6(1): 21–40.

[10] J. Drogalis, E. Keyes and L. Dhyani. "The Weather Channel application usability test", 2015. Retrieved from https://elkeyes.com/wp-content/uploads/2015/05/WeatherChannelUTResults.pdf

[11] International Organization for Standardization, Human-centered design for interactive systems (ISO 9241-210:2010), 2010.

[12] D. Norman. The design of everyday things: Revised and expanded edition. New York, NY: Basic Books, 2013.

[13] Y. Mao, K. Vredenburg, W. Smith and T. Carey. "The state of user-centered design practice". Communications of the ACM. 2005, 48(3): 105–109.

[14] R. Schnall, M. Rojas, S. Bakken, W. Brown, M. Carry and J. Travers. "A user-centered model for designing consumer mobile health (mHealth) applications (apps)". Journal of biomedical informatics. 2016, 60(2): 243–251.

[15] S. Jung, H. Kim, H. Lee. "Influences of perceived product innovation upon usage behavior for MMORPG, Product capability, technology capability, and user-centered design". Journal of Business Research. 2014, 67(10): 2171–2178.

[16] J. Nielsen. "10 usability heuristics for user interface design", 2020. Retrieved from https://www.nngroup.com/articles/ten-usability-heuristics/

[17] Y. Rogers, H. Sharp, and J. Preece. Interaction design: beyond human-computer interaction, Hoboken, NJ: John Wiley & Sons, 2011.

[18] J. Gove. "Principles of mobile app design: Engage users and drive conversions", 2022. Retrieved from https://www.thinkwithgoogle.com/future-of-marketing/creativity/principles-of-mobile-app-design-engage-users-and-drive-conversions/

[19] E. Boyatzis. Transforming qualitative information: Thematic analysis and code development, Thousand Oaks, CA: Sage Publications, 1998.

[20] V. Braun, V. Clarke. "Using thematic analysis in psychology". Qualitative Research in Psychology. 2006, 3(2): 77–101.

[21] J. Nielsen. "Severity ratings for usability problems", 1994. Retrieved from https://www.nngroup.com/articles/how-to-rate-the-severity-of-usability-problems/

[22] S. Yekta-Michael, C. Färber and A. Heinzel. "Evaluation of new endodontic tooth models in clinical education from the perspective of students and demonstrators". BMC Medical Education. 2021, 21(1): 447.

[23] J. Sauro, S. Dumas. "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, New York, 2009.

[24] P. Chin, A. Diehl and L. Norman. "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 system, New York, 1988.

[25] S. Joo. "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, Hoboken, 2010.

[26] E. Frøkjær, M. Hertzum and K. Hornbaek. "Measuring usability: are effectiveness, efficiency, and satisfaction really correlated", in Proceedings of the Special Interest Group on Computer-Human Interaction conference on human factors in computing system, 18th Annual Meeting, New York, 2000.