Jason Len and Sean Takafuji, Team Paper Prototyping, Team Assignment 1

Typical Tasks:

1. We want to test the system, Pretend that North Korea has launched a missile at Hawaii. What should you do?
2. This is not a test, North Korea has actually launched a missile attack against Hawaii. What should you do?
3. Opps, that last instruction was a mistake, what do you do now?

When presented with these tasks, the observer will present them in an ordered fashion. Each task will only be presented when the subject has indicated that they have completed the previous task. The tasks may be re-read to each participant as many times as requested. The attempt to not answer questions will be made, by repeating the prompt. If a question is needed to be answered by the observer, both the question and the answer given should be recorded. Order of slide presented was determined by what was clicked. The slide order is listed in Appendix B.

Metrics of Measurement:

In this assignment, it was decided that the metric of time while of significant importance in the real world, was not a metric that would be feasible in this environment. To measure this metric, each test user would have needed time to either be trained on the system or to have time to explore the system beforehand. In this scenario, most test users would not want to spend that amount of time to allow for the measurement of time to yield sufficient results.

Besides speed, it was also decided that accuracy was an important metric. In this case, this metric was testable. Towards this end, two aspects of accuracy would be measured, slips and mistakes. In a slip, the user would make a mistake that is due to them trying to do one thing, but accidently doing another. For example, if a user knew that they were attempting to send a test message out, but accidently clicked the wrong link and sent out a real warning message. In the metric of a mistake, the user knows what they were attempting to accomplish, but was not sure how to accomplish that task. An example of this would be where the user knew that they wanted to send out a test message, but actually sent out a real message because they did not realize that there was a difference between the corresponding links.

Unlike the first two metrics that will be measured, we will also attempt to measure the user experience of how comfortable were they in using this system. Since this metric will be more subjective, we will limit the answers to a scale of 1 to 10, where 1 will represent totally uncomfortable and 10 being most at ease. While this scale will limit the answer given, this metric will be supplemented by the additional open ended questions for a more in depth feedback. (See Appendix A for the full lists of questions.)

The above metric will be recorded in two parts. The first measurement will be taken immediately after the user has completed all tasks. At that point, the user will be given the opportunity to explore the system and answer the in-depth questions as mentioned above. At the end of the study, when the user is done and ready to leave, they will again be asked about their comfortability with the program. (It is believed/hoped that after being allowed to explore and after answering the question prompts, that their comfortability will increase.)

Data:
Subject #1 - Expert

Completed all three tasks with no error.
Initial comfortability: 5
Final comfortability: 6

This user found the interface to be fairly intuitive and “easy” to understand”. While this subject stated that they understood that the purpose of typing in their affirmative response to send the notice, was for redundancy, they did not agree with this on a philosophical level. They wanted a one click send to save time. “Even a few seconds matter.” When presented with the idea that by having the user also type in their answers was also to change the users mode of thought, the subject reiterated that they wanted a one click send methods as time was the most important factor to them. The subject did not have any opinions on the use of the pictures and the texts. But the subject did pose a question on the color scheme of red versus green. The subject speculated the idea that red could represent the actual missile crisis, so it should be used for the sending of the warning message, BUT the subject also believed that red could represent “stop” and should be used for the cancellation of the alert. The subject was not able to come to a conclusion on this color debate.

This subject stated that their comfortability was not based upon the usability, but instead, they really felt strongly about the one click send feature. This subject also added that they feelings were initially hampered by the fact that they needed to re-read the on-screen instructions a couple of times to ensure accuracy before they clicked. Because they did

this, they felt even stronger that they wanted a one click system, since they felt they would ensure accuracy before pressing their one button. They did add that they did see that if the personal that was using this system had adequate training, that they speed of sending out a message would be greatly reduced (as compared to their first experience), but it would never compare to their one click idea.

Subject #2 - Intermediate

Initially this subject executed the sequence to send out a real SMS text in scenario 1. After scenerio 2 was read, this subject realized that the might have made a mistake and wanted task 1 to be reread, upon hearing the instructions again, they realized their mistake and wanted to start again. On their second attempt, for the first task, they completed it correctly. The second and third tasks went smoothly.

Initial comfortability: 7.5
Final comfortability: 8.5

This subject was fine with pretty much everything, but wanted a greater difference between in the way the real button and test button looked. “A drastic change.” The subject did notice the difference between the pink and red, but would have preferred a more distinct color than pink. (Not a shade of red, but a truly distinct color such as purple or blue. This subject also noted that the picture on the test button was a clip art and that they picture on the real button was a real picture, but also wanted a greater change here. They suggested something link a nuclear explosion on the real button and a clown picture on the test button. Lastly, this subject thought that emphasizing the important word(s) would also be helpful.

Subject #3 - Intermediate
All task completed correctly.
Initial comfortability: 8
Final comfortability: 8.5

After completing the assigned tasks, this subject inquired about the other buttons that they did not press, and was told that they were free explore the system and press any button that they wished. In pressing the various buttons, this user did not find and asked about the inclusion of being able to send out a custom SMS text that could be entered “on demand”. They said that this would cover any unexpected situation that might occur. This user was also the only user that inquired about what the actual SMS message that this system would send would say. The subject was told that that information was outside the scope of this test scenario and had not yet been decided. This subject felt that the pictures, text, and spacing was good. In closing, this user also suggested that a confirmation screen of the text being cancelled would be good. (Not a confirmation to cancel, but a message confirming which alert was cancelled.)

Subject #4 - Novice
All tasks completed correctly.
Initial comfortability: 8
Final comfortability: 8

This subject felt that everything went “Okay”, but with the pressure of having a missile aimed at them, this subject worries that in the real situation, they might be too stressed to complete all the steps and that the steps might take a while. While this subject understood the need for verification (the typing of “send”), this subject would rather have had a simple click that needed a click verification for a different source (or two). This subject really liked the pictures and thought that the text added an extra level of security of understanding. But this subject did note that they did not really use the text to figure things out and only read the text to verify what they were already planning to do. This subject believes that the likelihood for error (in terms of both a mistake and/or slip was very unlikely. “Less than 0.5%.” Lastly, this subject asked why the button was called “Foreign Threat” and why it showed a picture of a missile. They wanted to see a more general picture. When told that a missile threat was the only current foreign threat, but it was labeled as such for future expansion, they said that they were satisfied with that answer.

Subject #5 - Novice
All tasks completed correctly
Initial comfortability: 8
Final comfortability: 9.5

This subject said that they found the tasks to be “difficult” but only because the did not want to read the text. They said that the largely ignored all of the text. In terms of the pictures, this subject wished for a different cancel button. When

prompted, they suggested a button with “cancelled” stamped over the button. This subject noted that they really liked the spacing between the buttons and actually was observed pressing the white space to see just how much room separated the buttons. They said they were satisfied with what appeared to be “more than enough space” between the buttons. This user believe that there was “no chance of a mistake” of either type happening if they user was trained on the system. Like Subject 1, this subject understood the purpose of typing send, but independently (both subjects were test outside of the proximity of each other) said that they had concerns about typing in the word “send” would take too long in a real emergency situation. When prompted about, “what if the user was adequately training on the system”, this subject still felt concerned. When prompted for their prefered method for verification, the subject suggested some type of verbal verification in which the user would say “Yes.” This user finished by stating that the felt that this system was very user friendly and easy to use.

Specific Improvements that may be implemented:

- Change the color scheme test button
- Change the picture of the test button
- Change the picture real missile attack button
- Add in a confirmation page for cancellation

Contributions of Members:

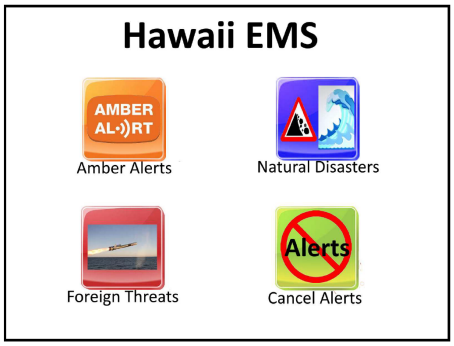
Jason - Base prototyping pictures and design
Sean - Edited and added in additional steps in prototype design
Sean - Designed scenario for testing and follow up questions
Jason - Wrote up testing procedures
Jason - Converted the pictures into PowerPoint clickable presentations for prototyping testing by users.
Sean - Conducted 3 tests
Jason - Conducted 2 tests
Jason - Wrote up results of all exams
Sean - Edited summary of results
Sean - Wrote up specific improvements
Jason - Edited specific improvements

Appendix A – Follow-Up Questions

· How easy or difficult was it to determine how to proceed in each scenario? Why?
· Which was easier to understand, the pictures or the text? Or were both necessary?
· Was it easy to understand what the pictures where trying to convey? Would you have been able to figure that out without the text? Which pictures do you think need improvement, and how?
· Was it easy to understand the text? Was there too much or too little text for each situation? Which texts need improvement and how?
· Did the colored backgrounds to each button help or hinder the process?
· Do you believe that after proper training user would be able to make a mistake on this system due to some type of confusion? (i.e. such as mistakenly thinking one button was another?)
· How likely do you believe it is that an error would happen due to an accidental slip such as a user unintentionally clicking or typing the wrong thing. NOT due to the lack of understanding? (i.e. Wanting to click on one thing but accidentally clicking the wrong item?)
· How easy was it to type in the correct answer to send the warning message?
· Why do you believe that the program would require the user to type in “send” instead of just clicking a button labeled send?
· How likely would it be for a person to accidently send a real message in a practice situation?
· How do you feel after looking over this program simulation?
· What went well?
· Was the system easy to use?
· Did you like it?
· What could be improved upon?

Appendix B – Slides and Order

Slide 1



Amber Alert links to Slide 10
Natural Disaster links to Slide 11
Foreign Threats links to Slide 2
Cancel Alerts links to Slide 8

Slide 2

Missile Attack Real links to Slide 3
Missile Attack Test links to Slide 12

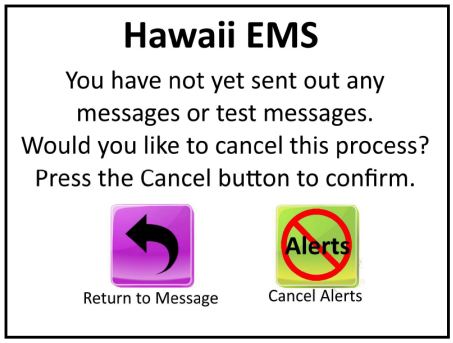
Slide 3

Enter links to Slide 4
Cancel links to Slide 6

Slide 4

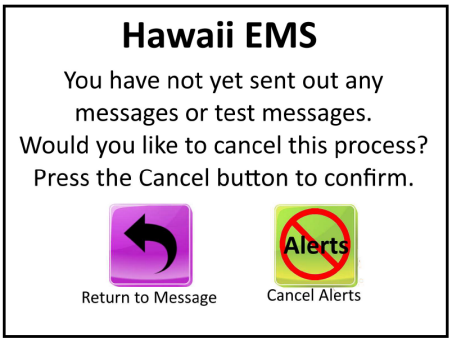
Main Menu links to Slide 1
Cancel links to Slide 8

Slide 5



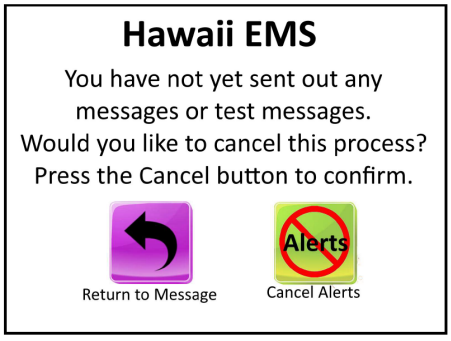
Return to Message links to Slide 2
Cancel Alerts links to Slide 3

Slide 6



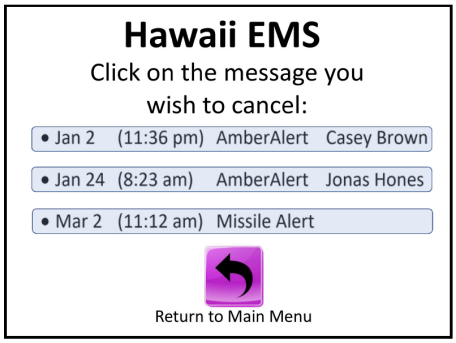
Return to Message links to Slide 3
Cancel Alerts links to Slide 1

Slide 7



Return to Message links to Slide12
Cancel Alerts links to Slide 1

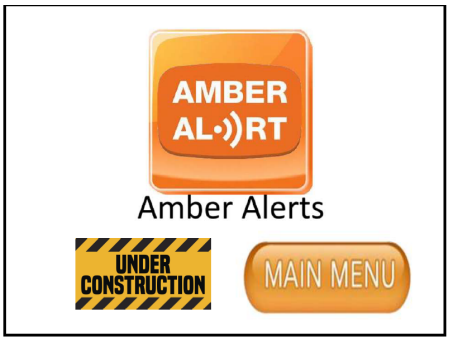
Slide 8



Both AmberAlerts link to Slide 10
Missile Alert links to Slide 9
Return to Main Menu links to
Slide 1

Slide 9

Main Menu links to Slide 1

Slide 10  Main Menu links to Slide 1	Slide 11  Main Menu links to Slide 1	Slide 12  Enter Links to Slide 13 Cancel links to Slide 7
Slide 13  Main Menu links to Slide 1		