
**SmartWatch Application
Report 2014**

HCI

Kenny Cox Reg. # 20549345 BU

Contents

Intro.....	3
Requirements of the Application.....	3
Methods used to evaluate the prototype.....	3
Quantitative	3
Issues that arose	3
Score/Rating Weight.....	4
Learn ability.....	5
Adapting to the user	5
Aesthetic Integrity.....	5
Feedback and Errors	5
Fulfilment	5
Conclusion.....	5
Qualitative.....	6
Heuristic Evaluation Results.....	6
Observation/Think aloud Results.....	6
Conclusion.....	7

HCI REPORT on SmartWatch Application

Intro

In this report, I will discuss the success and failure factors for my SmartWatch Application design by using practical heuristics and usability guidelines also, questionnaires were given to the three evaluators to fill out after completing a task list of requirements. I will discuss the findings from my quantitative and qualitative experiments, note the severity and frequency between issues and discuss the decisions behind my plan of action (based on severity+frequency). Graphs and charts used to identify faults further depict the status of my findings and my conclusion of my analysis. P. Zhang, and D. Galletta (2006) states "HCI seeks to maximize the use of information through the design of humanly acceptable representational and manipulatory tools."

Requirements of the Application

- Ability to synchronize the watch with your Uclan account. (Uclan logo)
- Ability to receive emails on the watch. (Uclan email connection)
- The timetable will be available to view. (Uclan timetable connection)
- It will have the ability to view a map of campus and assist in navigation around campus. (Uclan map)
- The ability to receive alerts such as last minute changes of rooms or staff sickness. (Uclan alerts) bell of some sort icon
- You may generate your own requirements based upon analysis of the various user groups.

Methods used to evaluate the prototype

Two novice evaluators and one expert evaluator between the age of 20-40. The evaluators tested the application on a HP laptop and used a mouse to control the application sitting down behind a desk and fresh water was provided. They were given instructions and given five (5) tasks (HTA, Hierarchical Task Analysis) to complete while testing the application. After they completed the tasks, they were given a series of usability, heuristic, user experience questions/statements and were asked to rate from 1-7 where 1 meant that they definitely not agree and seven (7) meant that they totally agreed. I observed the evaluators and took notes while they performed five(5) tasks that were on the requirement list as stated above in the "*Requirements of the Application*" section above. The three evaluators were given Jakob Nielsen "10 usability heuristics for user interface design" heuristics guidelines and D. Norman's "The Design of Everything" to compare and make a note of any design attributes that breached the guidelines. LOVE, S., 2005 states that "it is important that your evaluators are experienced mobile HCI practitioners or researchers".

Quantitative

The following are the problems and actions that need to be taken to improve the overall performance of the application. Shackel, (1986) identified four dimensions that are important: **effectiveness, learn ability, flexibility and attitude**. A severity rating was put in place where 1 is low, 2 is low-medium, 3 is medium-high and 4 is high. The results revealed that although my design seemed functional and obvious to the designer, experienced users found it difficult performing some tasks, as observed during the tasks and observation tests. Questions are represented as "Q" below, so "Q1" would be question 1 from the 20 questions/statements given. They are broken down into sections as shown below. The rating scheme is 1 for Bad/Defiantly not (agree) up to 7 which means Yes/totally agree and 4 would be the "undecided" number, and any rating under the number 4 is a failure and needs to be looked at and fixed (stated in the Problem, Severity, Frequency and Action table above). The most severe problems reported by the evaluators are listed below.

Issues that arose

The total merged problems found by all evaluators were not detrimental to the applications success, overall the

application did not have any dire faults, although there are some issues that arose during the evaluation. J. Annett, and N. Stanton (2000) stat that the "Hierarchical Task Analysis (HTA) developed by Annett et al. (1971) is probably the most frequently used methodology by ergonomists in the UK and is often incorporated into CTA, so it makes a useful point of reference when making comparisons with other approaches."

SEVERITY CHART																					
Problems listed in order of severity and frequency...problems that need to be modified first are on top																					
Problem				Severity Rating					Frequency					Action							
Ambiguity within the application is an issue. The "Campus" button leaves the user confused.				3					2					Rename the "Campus" button to "Uclan Map".							
The download button (on the alert pop up) is ambiguous to the user.				2					2					A pop up can notify the user what the icon is when he/she hovers over it, change icon to floppy disk icon.							
Alerts did not indicate the location, Needs to be more user friendly.				2					2					offer the user to access it without having to look for it.							
The design wasn't appealing to the eye.				1					1					A more cherry, pleasant on the eye colour scheme is required, maybe another evaluation on colour schemes is in order but for now I can use a three colour based scheme (blue, light blue, white, black(black is not considered a colour).							
The background light was not affective.				1					1					Enhance the brightness and offer alternative colours for the light source.							
Sound effects were not pleasant nor did they assist, boring was the verdict.				1					1					change the audio files to something more fun and user friendly.							
View appendix 1A for chart																					
Additional Actions required: Offer multiple colour themes to the users, Offer the users a "Help" button with application definitions, offer a downloads folder, use more metaphors in my design.																					
Practical Usability and User evaluation based in part on Nielsen's 1993 Heuristics and Norman's 1990 Principles, results below, charts are available in the appendix below.																					
LEARNABILITY					ADAPTING TO THE USER					AESTHETIC INTEGRITY					FEEDBACK/ERRORS			FULFILLMENT			
USER	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Q11	Q12	Q13	Q14	Q15	Q16	Q17	Q18	Q19	Q20	AVG
A	7	6	7	7	7	3	7	7	7	7	4	7	7	1	4	7	7	5	7	7	7.4
B	5	6	7	7	6	6	7	7	6	7	6	7	7	6	3	6	7	3	7	5	6.05
C	7	5	6	6	6	5	6	7	7	7	7	7	7	5	6	6	6	5	7	7	6.25
AVG	6.3	5.6	6.6	6.6	6.3	4.6	6.6	7	6.6	7	5.6	7	7	4	4.3	6.3	6.6	4.3	7	6.3	
Total AVG	6.275				6.35					5.58					5.73			6.65		6.117	
charts	View appendix 1B				View appendix 1C					View appendix 1D					View appendix 1E			View appendix 1F		View appendix 1G	

APPENDIX 1.K

Please View Appendix 1H, 1I, 1J for each of the evaluators results for Q1-Q20.

Score/Rating Weight

The rating scheme is as follows: 1 for Bad/Defiantly not agree, and 7 means Yes/totally agree. 4 being the middle of 1 and 7 would be the "undecided" number, and any rating with the rating score of 4 and lower is a failure. under the number 4 is a failure. In between the evaluators were completing their practical usability and user evaluation I held an informal interview to gather additional data, as to how they felt during their experience noted below. S. Love

(2005) stats that "Questioning an individual can be a quick way of getting information from them on usability issues, requirements gathering and a whole host of mobile HCI research interests."

Learn ability Q1 Evaluator B rated lowest which a rating of 5 isn't a bad one, it's still two score ratings lower than that compared to evaluator A and B's rating.

Evaluator A and B(both novice users) seemed to perform the same on questions 2-4. Q2 rated the lowest rating from all evaluators(dropped rating for each). With an overall all rating of 6.275, I'd say the learn ability features were successful.

Adapting to the user

Q6 rated the least in this category with all three evaluators, leaving me to conclude that the evaluators were not convinced that they were clearly instructed of what the limitations are within the application. I think that I pop up messages can be installed to better inform first time users on how the application works. Hints and friendly reminders would improve usability and over all the effectiveness and efficiency of the application. Other than Q6(rating of 4.6), the other 5 rated with a 6.3 or higher.

Aesthetic Integrity

A, B, and C all agreed that Q12 and Q13 both satisfied their needs in regards to buttons being large enough and easy to read elements. A gave the design a low rating for Q11, Q14, Q15 which lead me to conclude that they design was not appealing to his eyes, the background light needs to be improved (please see severity chart for plan of action or POA). Both A and B agreed that the colours (black and white predominantly) was not pleasant on the eye. colour scheme will be added and options for the user to choose different colour schemes. (please refer back to the POA for further details). Overall rating of a 5.58.

Feedback and Errors

Q18 rated very low with an average of 4.8 leaving me to conclude that the sounds were defiantly not pleasing to the ear and more work is required. The systems status and rollover effects were effective scoring a 6.3 and 6.6 respectively.

Fulfilment

Q19 and Q20 leaves me to conclude the application design features allowed the users to complete their tasks with minimum effort and strain. Successful in this area with an average rating of 6.65

Conclusion

A combined average rating of 6.117 out of seven (7). S. Love (2005)stats "The idea behind heuristic evaluation is that several people(known as evaluators) independently carry out a usability evaluation of a system or prototype to identify any potential problems with the design, using a known list of design heuristics as an aid." led to conclude that show that there are three major concerns and three minor concerns that need to be modified. By following a few general guidelines, it can correct for the majority of usability problems in advance. Although likert scale evaluation method isn't very conclusive, I chose to use it within my evaluation so that I can use to compare based on my other methodologies used. Patterns arose from my qualitative research and my thematic analysis left me to conclude that the application pros and cons were: **Pros:** functional, easy to use, fast, no learning curve, the system speaks the users language(for the most part, the design was consistent throughout, tailored for both experienced and inexperienced users. **Cons:** not appealing to the eye, ambiguous at some parts, feedback wasn't informative when saving(alerts), background light needs modifying/improving(not effective), sounds needs improving.

Qualitative

Heuristic Evaluation Results

Based on the problems reports in the heuristic evaluation, I found that the application does not keep to D. Norman's and Nielson's 1995 heuristics. The application failed in a couple places and was a success in others. Fail factors were based on heuristics values, I failed to offer "Visibility of System Status" which was a concern when the evaluators tried to download their alerts, it left them confused and limited their ability to access their data. Bryman, 2004; Cresswell, 2003, suggests that qualitative research is exploratory by nature and is used to explore phenomenon when the variables and theory base are unknown.

Observation/Think aloud Results

While observing evaluator A

Question 1 - the user did not require any assistance from me, they seemed to find it easy to navigate and use, user looked familiar with the program although he/she had never used it before, didn't struggle at all. found email easily.

Question 2 - found the email easy, didn't get lost, viewed both emails available by clicking on next arrow

Question 3 - yes all tasks were completed. very easy.

Question 4 - the system worked as expected but based on user interpretation or ambiguity within the compass's button name.

Question 5 - Male, between the age of 20-25

Question 6 -Smiling, seemed to enjoy his experience using the application. He said it was "straight forward & easy with no learning curve".

While observing evaluator B

Question 1 - Download/Save icon are ambiguous, Alerts doesn't show or keep the evaluator informed, the user didn't know where the alert was saved.

Question 2 - the evaluator could navigate easily, he said he was "impressed" and said that the battery for the home button was cleverly designed to replace the swipe function (since I did not test on a flat screen device).

Question 3 - yes the evaluator completed all his tasks, he started off sluggish but after completing the first two tasks the momentum sped up and he seemed to perform more comfortably and confidently.

Question 4 - the evaluator completed his tasks easily, without errors, he tried to break it, the system works as expected

Question 5 - Male, between the age of 30-35

Question 6 - the evaluator started off slow and gradually trusted and felt comfortable navigating the system after completing his first task.

While observing evaluator C

Question 1 - the evaluator did not require any assistance, he seemed to understand how to navigate the system as if he had used the application before. I can tell that he was very tech savvy.

Question 2 - the user clicked on each of the main buttons on the home screen, clicked the arrows within the secondary screens, he was familiar with the system before he read the task sheet, no learning curve.

Question 3 - all tasks were completed very quickly and easily

Question 4 - the system works as expected, successful and impressive.

Question 5 - Male, between the age of 30-35

Question 6 - The user felt comfortable and it looked like he was already with the application.

Conclusion

After reviewing the results from the observation results I think that a flat screen device and implementing the proper gestures in order to navigate any application will yield better results overall. Without observing, I would not have seen the difference in how the application was being used with first time users, it allowed me to see firsthand three experienced users/evaluators perform with some difficulty but for the most part, successfully.

Two out of three evaluators reported that they could not locate the Uclan map easily. This might have been due to the name of the button, it left the evaluator confused and had to second guess it. Also, the download button wasn't clear to the same two users, so if two out of three evaluators (66%) it is a major issue, it shows me that the novice users will have issues and experts won't have any issues. Evaluator C did not have any issues throughout the system, internet savvy people will not have any issues with this application. Since I did not have any females on my evaluation, I don't think the ratings are accurately transcribed because, I feel a woman's perspective is different than a male's perspective (at any internet/IT level). E. McKay, states that "According to Dr Brizendine, the female brain processes thoughts very differently compared with the way a male's does."

Overall I think the application was a success.

References

Bryman, 2004; Cresswell, 2003 - data gathered from your powerpoint

J. Annett, and N. Stanton, 2000, "Task Analysis",
London, GBR: CRC Press, 2000. p 2.
<http://site.ebrary.com/lib/uclan/Doc?id=10062976&ppg=9>

LOVE, S., 2005. *Understanding Mobile Human-Computer Interaction*. Jordan Hill, GBR: Butterworth-Heinemann.
<http://site.ebrary.com/lib/uclan/Doc?id=10138711&ppg=79>
Accessed on March 30, 2014

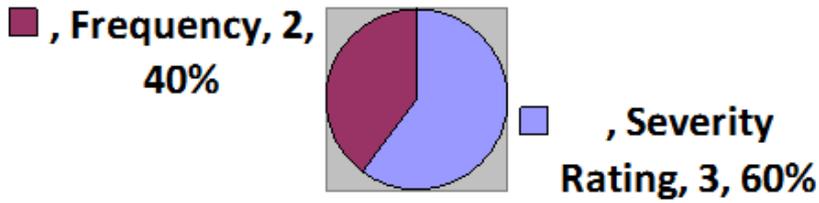
McKay, E.. *Human-Dimensions of Human-Computer Interaction*.
Amsterdam, NLD: IOS Press, 2008. p 202.
<http://site.ebrary.com/lib/uclan/Doc?id=10267478&ppg=222>

Shackel, (1986) - data gathered from your powerpoint

P. Zhang, and D. Galletta, 2006, *Human-Computer Interaction and Management Information Systems : Foundations*.
<http://site.ebrary.com/lib/uclan/Doc?id=10178097&ppg=21>

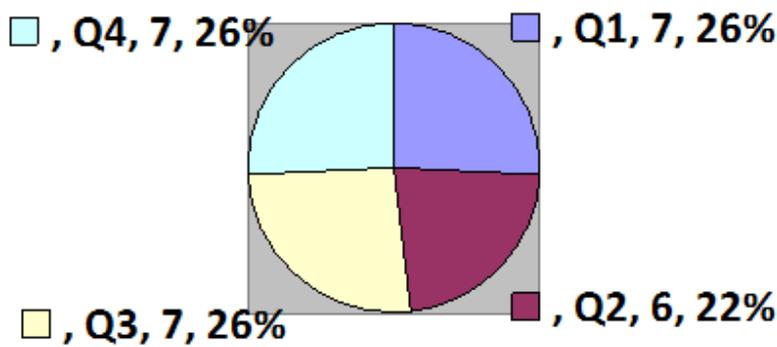
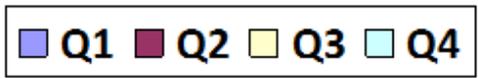
1.A

Severity Rating Frequency



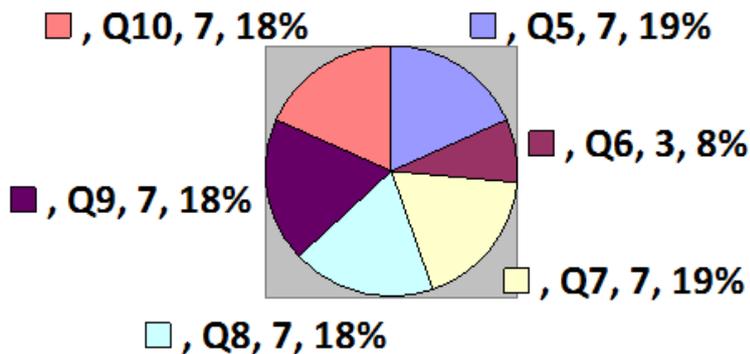
1.B

LEARNABILITY



1.C

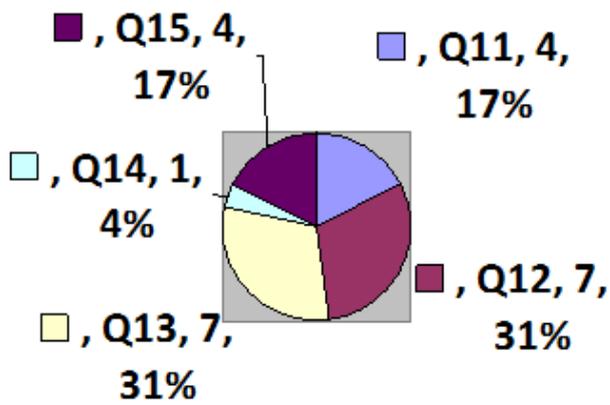
Adapt to the user



1.D

AESTHETIC INTEGRITY

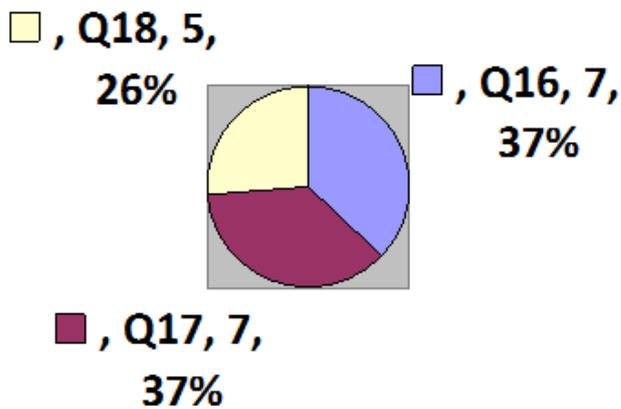
Q11 Q12 Q13 Q14 Q15



1.E

FEEDBACK/ERROS

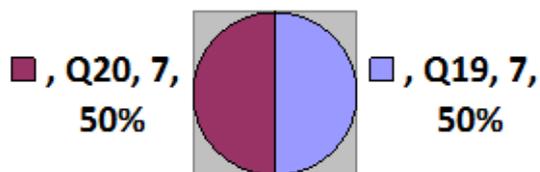
Q16 Q17 Q18



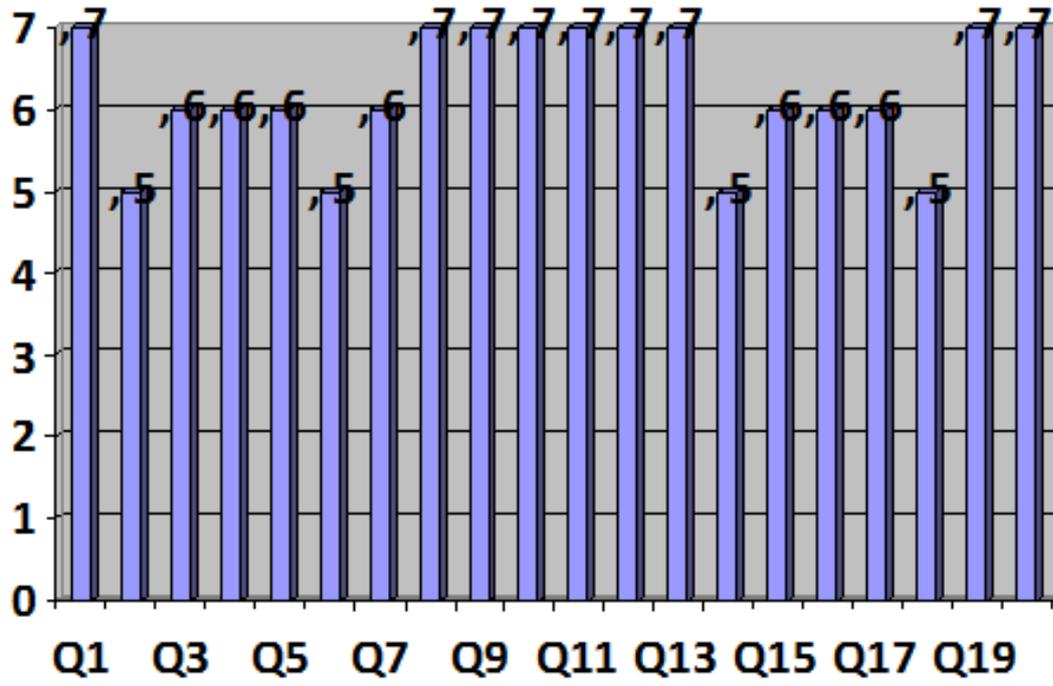
1.F

FULLFILLMENT

Q19 Q20



1.J



1.K

