Designing quick and dirty applications for mobiles: making the case for the utility of HCI principles

Baillie, Lynne; Morton, Lee

Published in:
Journal of Computing and Information Technology

DOI:
10.2498/cit.1001810

Publication date:
2010

Document Version
Publisher's PDF, also known as Version of record

Journal of Computing and Information Technology - CIT 18, 2010, 2, 103–107
doi:10.2498/cit.1001810

Designing Quick & Dirty Applications for Mobiles: Making the Case for the Utility of HCI Principles

Lynne Baillie and Lee Morton

Glasgow Caledonian University, School of Engineering & Computing, Cowcaddens Road, Glasgow, Scotland, UK

Many applications are currently being built for mobile phones that are intended as throwaway gimmicks that people download from places like Apple Appstore. Users can download small throwaway applications for their mobile phone for as little as ninety nine cents. We were interested in what effect these two components e.g. throwaway and cheapness have on the use of HCI guidelines by the designers of these applications and whether or not it was worth their while incorporating them into their design given the temporary nature of use. In this paper we describe how we tested two designs of the same concept. The first design brief was company-led and did not explicitly adhere to any HCI principles and the second was designed according to HCI principles. We tested both applications with users in the field to see which was the simplest and most intuitive to use.

Keywords: mobile application, throwaway application, usability, evaluation

1. Introduction

This is a situation not unfamiliar to HCI practitioners to be asked to justify the costs of usability testing and of adhering to HCI design principles when a lot of applications, systems and products have been built without adherence to them[7][6]. Our contention, and one which we believe is echoed by [2], is that not adhering to them often comes with a cost and that our artifacts might have worked better, they might not have failed, they might have been more user friendly. Our hypothesis therefore was that: (HA) not adhering to HCI design principles would have an important impact on the user experience of a throwaway application.

That is why we set up an experiment in which two designers with similar educational backgrounds (e.g. masters in computing) and work experience (short length of time e.g. less than 6 months designing applications) designed a mobile chill application. They were given the same brief and the same time in which to build the mobile application.

In this paper we present the concept and evaluation of the two applications and draw conclusions from the results as to whether the outcome proved the above hypothesis.

2. System Overview

The concept is that of a mobile drinks timer that you can set to chill your drinks. The Main Menu should offer the user three different options:

- "Start New Timer", this function should display a series of menus allowing the user to select the required chill settings for their drink and start a timer.
- "Start Saved Timers". The start saved timers option displays a list of previous timers that have been used in the past and that the user can use again. When the timer finishes, a tone is played to alert the user.
- "Preferences" this function should allow a user to set their own preferences e.g. type of alert to be sounded once drink is chilled.

The HCI driven application was designed using five main guiding design principles: constraints, consistency, affordance, visibility and feedback. One of the designers worked from the company design brief which contained no mention of HCI principles and designer B from the design brief which, in addition to the company brief, also contained HCI principles and guidelines. In the paper, we refer to them as designer A and designer B.

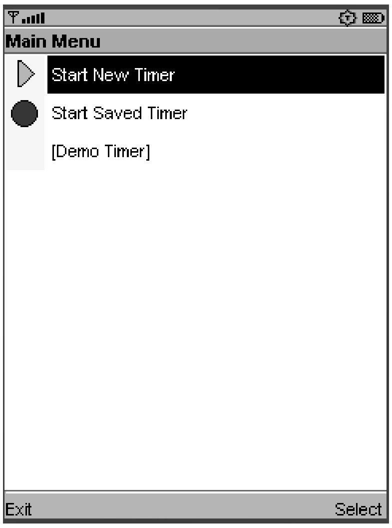

Figure 1. Main menu of application produced by designer A.

Figure 2. Main menu of application produced by designer B.

3. Evaluation

We had a mix of participants who were recruited from a sports club. None of the participants were university or company employees, but all owned mobile phones. Ten participants (six male, four female) took part and they were aged between 21 and 56 with an average age of 29. We undertook the testing in the participants own home for two reasons: one because this is where the application would be used and because we

Figure 3. Application running on various devices (Nokia 5500 left, Sony Ericsson P990 middle, Blackberry K800 right).

[1] have found that testing mobile phone applications in a lab environment can be problematic as it does not provide a good indicator of actual use.

Three different devices were chosen for testing (see Figure 3). The Nokia and Sony-Ericsson were used as examples of a standard mobile phone and a Blackberry as an example of a touch screen PDA type phone.

4. Methodology

Most empirical evaluations of input devices or interaction techniques are comparative. A new device or technique is compared against alternative devices or techniques. One design for such experiments is the repeated-measures design, also known as a within-subjects design [4]. In a repeated-measures design, each participant is tested under each condition. The conditions are, for example, device "A" and device "B". So, for each participant, the measurements under one condition are "repeated" on the other conditions. There are drawbacks to this method though and as a result it cannot be used on its own. The drawback is called "learning effects", known more formally as "asymmetric skill transfer"[8]. To compensate for this, a technique known as counterbalancing was used. Typically, counterbalancing is performed by placing participants in groups and presenting conditions to each group in a different order. As a result, half of our users were asked to complete task set 1 on phone A then set 2 on phone B, the other half completed set 1 on phone B then set 2 on phone A. This is to attempt to see if there is a significant difference in the usability of the application on different devices.

The tasks given were not designed to be step by step instructions, but were designed to emulate real world tasks and test whether a user would know the correct action to take from various screens. Each task was timed and given a score based on the user's ability to complete the task successfully. The scoring system is similar to that used by usability experts[5], where each score represents the success level of completion rather than the severity of the error.

To gauge the users' opinion after completing the tasks they were asked to complete a System Usability Scale questionnaire similar to that devised by [3] to assess their opinion of the software. This scale cannot directly measure usability, but can be useful in comparing results from different users as well as providing a general positive or negative feedback on the users' satisfaction with an application.

Before completing the tasks, the users were asked to complete a quick questionnaire detailing their experience with mobile phone applications.

5. Results

Not all the features and functions of the two application match as they were designed by two different designers. However, one feature that was identical is the process of starting timers.

Task	Description
1	Start a timer for a 330 ml can, in a fridge, lightly chilled.
2	*Switch off the vibrate function.
3	Delete the timer you just started.
4	Start a timer for a 500 ml bottle, in a freezer, extra cold.
5	Exit the application.
6	*Select "tequila" as the alarm sound.
7	Start a timer for a wine bottle, in a freezer, extra cold.
8	Start a timer for a 330 ml bottle, in a fridge, lightly chilled.
9	*Reduce the volume level of the alarm sound.
10	Delete the timer for the wine bottle.

Table 1. Shows the tasks undertaken (*shows what tasks could not be completed by users using application designed by designer A).

Tasks 1, 4 and 6 could not be completed by any of the users at the first attempt using Designer A's application.

Users were able to complete each given task in an average of 8 seconds. The average time taken to set a timer was 6.9 seconds excluding the first attempt (to allow for learning). These times are suitably low for a mobile device as researchers [1] have found that once a task takes more than 30 seconds on a mobile device, users start to

Task	Design A		Design B	
	Time *(s)*	Score *(1-4)*	Time *(s)*	Score *(1-4)*
1	16.4	3.4	9.4	4
2	14.4	3.4	10.6	3.8
3	6.4	3.8	8.4	3.6
4	7.2	4	6.4	3.8
5	8.8	3.8	5	3.6
6	7.8	3.4	6.4	4
7	8.4	4	5.2	4
8	8	3.8	6.4	4
9	7.8	3	8.2	3.6
10	6	3.4	4.6	3.4
Average	**9.12**	**3.6**	**7.06**	**3.78**

Figure 4. Average task completion times and scores arranged by application.

become frustrated. We can conclude therefore that the time factor is not likely to discourage people from using either of the applications.

Some users were irritated by the multiple button pressing required to start the timers in design A and this was responsible for some of the errors as users hit the save button before having the appropriate menu item highlighted.

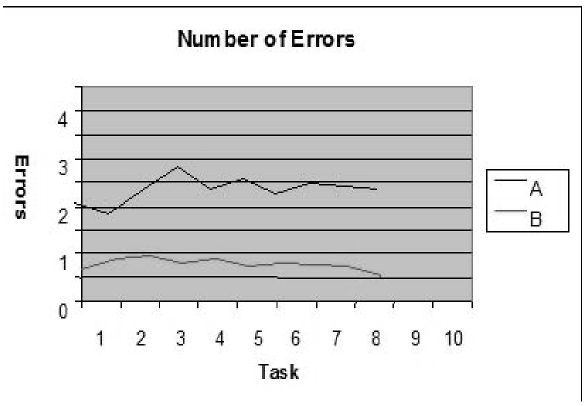

Figure 5. Number of errors on average each user incurred using the application during each task.

Users also took longer to learn how to operate design A. They generally completed two or three tasks with errors before becoming proficient whereas the majority of users of application B were proficient immediately or after one attempt.

We used the System Usability Scale (SUS) to measure user satisfaction with the applications. This satisfaction scale has been around for several years, and is used by many usability testers, organisations and websites e.g. Intel, Wikipedia and NASA. The SUS scale uses 10 categories to evaluate satisfaction. SUS scores can range from 0 (very little satisfaction) to 100 (very high satisfaction). Average satisfaction scores are usually between 65 and 70. Users rated the application B at an average of 74 out of 100 on the system usability scale (SUS) with a maximum of 92 and minimum of 60. This indicates that the users, on the whole, rated the application highly and found it easy to use. Application A was not rated so highly with an average of 45 with a high of 62 and a low of 37.

We found that there is a direct correlation between the users' previous phone experience and the score reported on the SUS. Those with more experience tended to report higher scores on the SUS. This may be due in part to the application emulating standard user interface components, as recommended by other researchers [9], found on phones.

6. Conclusion

This paper reported on the design and testing of two throwaway mobile chill applications. By their very nature these applications must be quick to design and build and as a result the designers were given the same time to design each application. The interface produced by B was quicker and easier to learn and use, and produced less errors. Users were able to complete each given task in an average of 18 seconds. This time is suitably low for a mobile device as researchers [1] have found that once a task takes more than 30 seconds on a mobile device, users start to become frustrated. We can conclude therefore that the time factor is not likely to discourage people from using either of the applications.

We have therefore gone some way to proving our hypothesis set out in our introduction i.e. that this application worked better, did not fail as often and was more user friendly because one of the designers was given HCI guidelines to follow and implement.

There are of course the usual caveats to attach to our findings e.g. the small evaluation conducted, the inherent talent of a designer and the type of mobile application itself. However, we still believe that, were the experiment to be repeated, we would anticipate seeing very similar results. The reason for this is that we focused on the tasks to be completed and assessed this rather than the aesthetics of the design.

7. Acknowledgments

We would like to thank Visual Products and ED Technology Support Centre for Scotland for supporting this work.

References

[1] L. BAILLIE, R. SCHATZ, Exploring Multimodality in the Laboratory and the Field. *Proceedings of the 7th International Conference on Multimodal Interfaces*, Torento, Italy, 2005.

[2] E. BIJKER, J. LAW, Shaping Technology and Building Society. Massachusetts, MIT Press, 1992.

[3] J. BROOKE, SUS – A Quick and Dirty Usability Scale, Usability evaluation in industry, Chapter 21. Published by CRC Press, 1996.

[4] A. DIX, J. FINLAY, G. ABOWD, R. BEALE, Human-computer Interaction. 2nd edition, Prentical Hall, 2004.

[5] KARAT ET AL, Comparison of Empirical Testing and Walkthrough Methods in User Interface Evaluation. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, Monterey, California, United States, pp. 397–404, 1992.

[6] G. LINDGAARD, Introducing HCI into an Organization: Making a Convincing Case for Usability. *Proceedings of INTERACT*, Rio De Janeiro, Brasil, 2007.

[7] J. NIELSEN, Guerrilla HCI: Using Discount Usability Engineering to Penetrate the Intimidation Barrier. *Cost-justifying usability*, Pages: 245–272, Academic Press, Inc., 1994.

[8] E. C. POULTON, , P. R. FREEMAN, B. N. MUKHERJEE, S. M. PITTEL, Unwanted asymmetrical transfer effects with balanced experimental designs. *Psychological Bulletin Journal*, Volume 66, Issue 1, 1966.orghoff V, Paresi R, editors. Information Technology for Knowledge Management. New York, Springer Verlag, 1998.

[9] M. F. QUSAY, Learning Wireless Java, Chapter 5 published by O'Reilly & Associates, 2002.

Received: June, 2009
Accepted: April, 2010

Contact addresses:
Lynne Baillie
Glasgow Caledonian University
School of Engineering & Computing
Cowcaddens Road, Glasgow
G4 0BA, Scotland, UK
e-mail: l.baillie@gcal.ac.uk

Lee Morton
Glasgow Caledonian University
School of Engineering & Computing
Cowcaddens Road, Glasgow
G4 0BA, Scotland, UK
e-mail: lee.morton@gcal.ac.uk

LYNNE BAILLIE, is a Reader and Director of the Multimodal Interaction Group at Glasgow Caledonian University, UK. She first attended ITI in 2002 and has been back four times since this initial visit. Once she even drove from Vienna all the way there (this was before the motorway was built), such was her enthusiasm to attend! She has published over 50 peer-reviewed scientific papers in the areas of Human Computer Interaction (HCI), pervasive computing, mobile and home environments. She believes that the reason that she returns to the ITI conference is that the people who attend have open minds and hearts and that this is a very safe and positive environment in which to discuss her groups latest research.

LEE MORTON is a Research Associate working in the Multimodal Interaction Research Group at Glasgow Caledonian University, UK. His research interests are: novel interaction modalities, interactive navigation, novel interface solutions for particular user groups and concepts on how to integrate mobile digital media with physical actions. He holds a postgraduate diploma in Advanced Computing and Masters Degree in physics.