A Mixed Method for Evaluating Input Devices with Older Persons

Murni Mahmud
School of Informatics
The University of Manchester
PO Box 88, Manchester M60 1QD
+44 161 306 3337
m.mahmud-2@postgrad.manchester.ac.uk

ABSTRACT
This research is an exploratory study which introduces a mixed method in evaluating common input devices. The method includes both quantitative and qualitative approaches and considers both subjective and objective measures. The study incorporates psychometric tests to measure user ability, introduces real tasks in the evaluation, and interviews users to elicit their opinions regarding the important qualities of preferred devices. A mouse, a tablet-with-stylus and a touch screen have been evaluated in two tasks: browsing a website and playing a card game. This paper shows the mixed method has made possible a more nuanced understanding of the use of input devices by older persons.

Categories and Subject Descriptors
B.4.2 Input/Output Devices; H.1.2 User/Machine Systems (e.g. Human Factor).

General Terms
Performance, Experimentation, Human Factors.

Keywords
Evaluation Method, Experimental Design, Input Device.

1. INPUT DEVICES EVALUATION
An input device is an intermediary which enables the interactivity between user and computer interface. Users' task performance when using input devices may depend on several factors: their motor speed (modelled by Fitts' Law [5]), cognitive abilities (some devices are more complex to operate than others, causing high cognitive load which might affect task performance), age (research on ageing shows that older people are slower than their younger counterparts [1] due to ability decline) and the task itself (for example, while a touch-screen might be the fastest and most errorless device for sketching, it might not be the best device for high-resolution drawing [8]). In addition, the most preferred input devices will also depend on the tasks at hand.

Many input device studies seem to consist of experimental rather than real tasks, partly because ISO 9421-11[9] recommends Fitts' Law experimental task as a standard evaluation procedure for input devices. While experimental tasks allow controlled manipulations of stimuli, necessary for performance measurement, the results might not be generalisable to actual life situations as in the cases where real tasks were used.

Older people seem to be particularly at a disadvantage when using common input devices. For example, using a mouse places four burdens on many older people: the need for precision; quick movement; coordination of shoulder muscles; and a high degree of eye-hand coordination [4]. Other studies also point out that some older users have negative perception toward computers. Not many studies, however, have combined several methods to get a more flexible picture of the use of input devices by older persons. This research introduces a combination of qualitative and quantitative as well as subjective and objective methods, with the main aim of getting a holistic view of this issue; it was run with 12 subjects.

2. A MIXED EVALUATION METHOD
The mixed method integrates four methods: (1) testing user ability with psychometric test, (2) capturing subjective ratings with input devices at pre- and post-tasks, (3) performing two real tasks with three selected input devices, and (4) eliciting user opinions regarding the qualities of input devices that they think are helping or hindering their tasks.

Conducting the mixed method has several advantages: (1) it is practical to conduct and provides a vast amount of data, (2) it provides an amusing and realistic environment, (3) no specific certification is required to conduct the psychometric test, and (4) it is highly portable.

2.1 Experimental Procedure
An experiment was conducted to investigate the effectiveness of the mixed method, according to the following procedures:

- Demographic questionnaire to elicit subject background, experience using computers, the internet and input devices.
- Ability measures with established psychometric tests: Mini Mental State Examination (MMSE – for cognitive abilities), perceptual with Identical Picture test (IP), and motor performance with Simple Reaction Time (SRT).
- Performance using input devices for two real tasks: web browsing (point and select) and playing a card game (point, drag and drop). We argue that we cannot test systems with only one task as there is a high risk of task bias in the results.

ASSETS'06, October 22–25, 2006, Portland, Oregon, USA.
ACM 1-59593-290-9/06/0010.

- Ease-of-use subjective ratings at pre-task and post-task. We argue that the task performance might influence user perception, and therefore we need to record the subjective opinions before and after the task sessions.
- Justifications for ratings and experience of using the devices.

The selection of psychometric tests was based on the simplicity, time and cost of conducting the tests. The selected input devices were a mouse, a tablet-with-stylus and a touch screen. They were selected based on their availability and guidance from [2]. Web browsing and the card game were selected to represent the task environment as they are two of the most common tasks older persons use computers for. Further description of the experiment and partial results are available in [6] and [7].

3. RESULTS AND DISCUSSION

The following results illustrate some examples of how the mixed method provided a more nuanced understanding of the use of input devices by older persons. They could be used as guidelines for selecting appropriate input devices that fit the users' general ability and task at hand.

In experimental tasks the mouse was often found to be the most efficient device. What our experiment showed is that while in the card game the mouse was the most efficient device, the touch screen was the most efficient one for browsing. Without using two real tasks, the fact that device performance was affected by the nature and complexity of the tasks would not have surfaced.

Recording user opinions pre- and post-tasks had allowed us to observe that users were more objective and realistic after they performed the required tasks. The post-task ratings were more closely related to their objective performance than the pre-task ratings. That is, the mouse received a more optimistic rating and the tablet-with-stylus and touch screen received less optimistic ratings in term of their ease of use.

We started the research by incorporating the psychometric tests to measure users' abilities and to understand the relationship to users' performance. MMSE was a very good instrument for checking that the participants were cognitively stable. However, there was not enough variance between users to make MMSE a meaningful way to measure users' cognitive variations The IP was a good indicator of users' performance, while the SRT was a good indicator of users' experience with computers, internet and input devices. Without running these tests, this unique relationship would not have been revealed.

Finally, we needed to know why certain input devices were considered user-friendly and usable but others were not. Users' justifications were elicited through debriefing interviews, which were analysed and categorised using a content analysis. Figure 1 shows the three important themes: *Frequent themes* were drawn from the number of occurrences of the word in the justifications, *Recount themes* were the key subject matter being discussed in the justification, and *Quality Themes* were the essential characteristics of input devices derived from the literature. These themes uphold the definition of usability and accessibility [9] and agree with the importance of ergonomic and human factors in evaluation of input devices.

The attitude toward each quality of an input device provides an explanation of user preference. Users have more positive attitude toward a mouse because a mouse has all of the quality themes; the strongest are effectiveness, ease of use and familiarity. User attitude was adverse toward a touch screen because it has low effectiveness, physical environment and efficiency even though it is superior for accessibility. For tablet-with-stylus, users have adverse attitudes because of low effectiveness, satisfaction, physical environment and efficiency.

The mixed method has been used widely in other research fields [2]. This research has shown some evidence that using a mixed method of evaluation has allowed us to understand why certain devices would be preferred by older persons and others would excel in certain types of tasks. If we had run only some of the experiments, we would have missed out the bigger picture. One might argue that the whole experiment sounds meticulous and lengthy. However, the whole session only lasted for one hour, and the effort was worthwhile as it produced more holistic knowledge regarding the use of input devices by older persons. The studies reported in [6] and [7] provide evidence that this mixed method is robust, and can be directly applied to other studies on interactive systems and older persons. Another domain application in which this method may be useful is the area of ubiquitous computing, which may involve multiple devices and tasks.

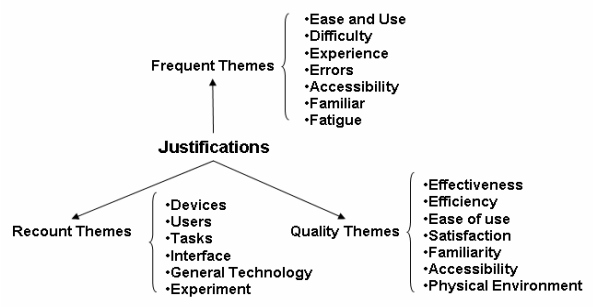

Figure 1. Frequent, Recount and Quality Themes.

REFERENCES

[1] Czaja, S. J., Sharit, J., Ownby, R., Roth, D. L. & Nair, S. Examining Age Differences in Performance of a Complex Information Search and Retrieval Task. *Psychology and Aging* 16(4) (2001), 564–579.

[2] Creswell, J. W. Research Design: Qualitative, Quantitative, and Mixed Methods Approaches (2nd ed.). Sage Publication, 2002.

[3] Greenstein, J. S. & Arnaut, L. Y. Human factors aspects of manual computer input devices. In G. Salvendy (Eds), *Handbook of Human Factors*. New York: Wiley, 1987, 1450-1489.

[4] Laursen, B., Jensen, B. & Ratkevicius, A.. Performance and muscle activity during computer mouse tasks in young and elderly adults. *Eur J Appl Physiol*., 84(4) (2001), 329-336.

[5] MacKenzie, I. S. Motor behaviour models for human-computer interaction. In J. M. Carroll (Ed), *Toward a multidisciplinary science of human-computer interaction*. San Francisco: Morgan Kaufmann, 2003, 27-54.

[6] Mahmud, M. & Kurniawan, S. Incorporating User Ability Test in Evaluating Pointing Devices. In *Proceedings of HCI International, Las Vegas*, 2005 (CD-ROM).

[7] Mahmud, M. & Kurniawan, S. Involving psychometric tests for input device evaluation with older people. In *Proceedings of the 19th OZCHI,* 2005.

[8] Milner, N. P. (1988). A Review of Human Performance and Preferences with Different Input Devices to Computer Systems. In D. Jones & R. Winder, (Eds.), *People and Computers IV: Proceedings of HCI'88,* Cambridge University Press, 1988, 341-362.

[9] Smith, W. J. *ISO and ANSI ergonomic standards for computer products : a guide to implementation and compliance*. Upper Saddle River, NJ: Prentice Hall, 1996.