Using Paper Mockups for
Evaluating Soft Keyboard Layouts

I. Scott MacKenzie[1] and Janet C. Read[2]

[1]Dept. of Computer Science & Engineering
York University
Toronto Canada M3J 1P3
mack@cse.yorku.ca

[2]Dept. of Computing
University of Central Lancashire
Preston, UK PR1 2HE
jcread@uclan.ac.uk

Abstract

Five experiments were conducted to compare soft keyboard layouts. The methodology involved paper mockups and manual timing in a classroom situation. Students worked in pairs – one as experimenter, one as participant – swapping roles midway through the experiment. Participants used a stylus to tap the well-known "quick brown fox" phrase five times on each layout. Entry speeds, computed from the measured time to enter the phrase, were 26.5 to 34.5 wpm for the Qwerty keyboard layout, 12.3 to 14.7 wpm for the Opti layout, 15.7 wpm for the Fitaly layout, 12.1 and 12.3 for a Qwerty-Phone (QP) hybrid layout, and 19.0 to 23.0 wpm for the standard phone keypad layout. The merits and limitations of the evaluation method are discussed.

Keywords

H.5.2 [Information Interfaces and Presentation]: User Interfaces, soft keyboards, text entry, empirical research methods, walk-up usability, paper mockups

1 Introduction

Developing efficient methods of text entry is a popular research topic in today's race for new mobile communications products. Given the ever-shortening time to market of new initiatives, developing efficient methods of evaluation is a desirable adjunct to research. This paper is primarily concerned with the latter of these two themes – to formally develop, test, and critique a

rapid evaluation method for new input techniques. The problem is presented in the context of the former theme – the development of efficient means of text entry for mobile systems.

1.1 Mobile Text Entry

Despite the appeal of miniaturization, mobility bears a price. The physical means for input are constrained by the small form factor, and, so, full-size keyboards and mice are not practical. Other input mechanisms are required, such as speech, physical keyboards with fewer or smaller keys, or stylus input.

Stylus-based mobile systems typically support two forms of input: gesture recognition and tapping. Stylus tapping on a graphic representation of a keyboard – a soft keyboard – is popular for text entry, as its use is common on stylus-based mobile devices. A soft keyboard is easy to implement and provides an alternative to handwriting.

With physical layouts, layouts other than Qwerty and phone are of little interest today. Alternatives, such as Dvorak [20], alphabetic [23, 33], or chord keyboards [7, 17], can support higher entry rates; however, substantial practice is required to gain proficiency. This, combined with a large installed base for Qwerty, has ensured the continued role of Qwerty as the keyboard of choice for desktop computing.

For soft keyboards, the arguments for Qwerty are diminished. Since the device is virtual, rather than physical, manufacturing costs lie in the software, and are one-time only. Thus, exploring the design space of soft keyboard layouts has emerged as a significant area of research [6, 12-14, 18, 19, 26, 28, 29, 34, 39, 41-43].

1.2 Expert vs. Novice Users

Most work on the design of text input methods focuses on the potential or expert entry rate of a

design [12, 19, 26, 42]. However, the novice experience is paramount for the success of new text input methods [25, 27]. This is at least partially due to the target market. Mobile devices, such as cell phones and PDAs, once specialized tools for professionals, are today used by consumers. It follows that immediate or walk-up usability is important. In other words, it is a moot point to establish the expert text entry rate if prolonged practice is required to achieve it. Consumers, discouraged by their initial experience and frustration, may never invest the required effort to become experts.

1.3 Evaluation

Empirical evaluations of new interaction techniques are time consuming and labour-intensive. And so, a related research topic is the development of efficient methods of evaluation. There are a variety of such methods in use, such as "wizard of oz", where the user unwittingly interacts with a human instead of a system [2, 4, 5, 11, 16, 30, 31]. Clearly this is efficient, since implementation is delayed until evidence is gathered on problems in the interface.

Paper mockups provide another convenient and efficient means to gather feedback from users. In this case, an interface is implemented on paper and user impressions are solicited, perhaps across several hypothetical implementations. While such methods are popular and successful, prior work with paper mockups is exclusively qualitative [1, 3, 8-10, 21, 32, 35, 37, 38]. Our interest is to explore the use of paper mockups for quantitative evaluations. This is important here since the most common research questions on text entry pertain to entry speed.

To increase the efficiency of the evaluation, our method involves the simultaneous testing of all participants, and engages the participants as experimental assistants in the evaluation. Our methods, results, and analyses were tested in five experiments.

2 Method – Experiment 1

2.1 Participants

Twelve university-aged volunteer participants were recruited from the local university. All were students enrolled in a course in human-computer interaction. The participants also served as assistants in conducting the experiment (see Procedure below).

2.2 Apparatus

The Qwerty and Opti soft keyboards layouts were selected for evaluation. Opti is a high performance layout [26] designed using the Fitts-digraph model of Soukoreff and MacKenzie [39]. The predicted expert entry rates are 30.0 wpm for Qwerty and 42.2 wpm for Opti [24]. Both layouts were implemented as paper mockups. See Figure 1.
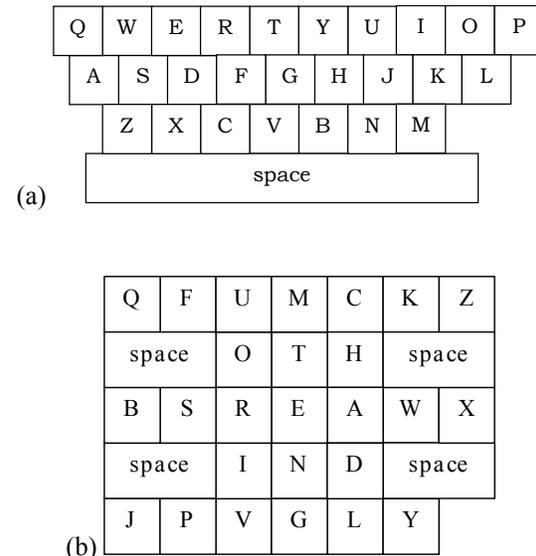

(a)

(b)

Figure 1. Soft keyboard layouts used in Experiment #1. (a) Qwerty (b) Opti

As measured on the paper mockup, the Qwerty layout was 9.6 by 3.6 cm and the Opti layout 7 x 5 cm. These dimensions are larger than typical for soft keyboards on PDAs. However, this should not impact performance, as there is both theoretical and empirical evidence [27] that, within reason, text entry rates for soft keyboards are not affected by the size of the layout.

With this highly-simplified apparatus, entry times could not be electronically measured, as there was no sensing technology or experimental software. Entry times were hand recorded with a timing device, such as a sport watch or a mobile phone in timing mode.

2.3 Procedure

Participants were instructed to study and memorize the following 43-character phrase:[1]

```
the quick brown fox jumps
over the lazy dog
```

The phrase was entered by tapping on the soft keyboard layout with a stylus. Participants provided their own stylus. Most used either a pen with the tip covered (or held upside down), or a pencil with the lead retracted. Used in this manner, the layout sheet remained clear of marks throughout the testing.

The instructions were to enter the phrase "as quickly as possible while trying not to make mistakes". Since no text was generated and accuracy was not recorded, some additional clarification was given on the need to proceed quickly (not recklessly) while accurately tapping the correct keys on the soft keyboard.

The participants worked in groups of two: one tapped while the other timed. A trial began when the timer said "start". Since no text was generated electronically, it was difficult for the timer to follow the progress of input. And so, participants were instructed to say "stop" upon tapping the last character (the "g" in "dog"). Timing was thus terminated for the phrase. The measurement in seconds was entered in a log sheet.

The procedure above was repeated five times using one layout, then five times using the other layout. Following this, the participants reversed their tapping and timing rolls and repeated the procedure.

To compensate for learning effects due to the order of testing layouts, participants were divided into two groups. Six participants entered with the Qwerty layout first, followed by Opti. The other half reversed the order. The experiment was conducted in a classroom as part of a regularly scheduled lecture for a course in human-computer interaction. The total time to conduct the experiment was about 40 minutes.

2.4 Design

The experiment was treated as 2 x 2 x 5 mixed design. Group was a between-subjects factor with two levels (Group 1 vs. Group 2, six participants per group). The within-subject factors were Layout with two levels (Qwerty vs. Opti) and Trial with five levels (1, 2, 3, 4, 5). The total amount of input was 6 participants/group x 2 groups x 2 layouts x 5 trials = 120 phrases.

Entry time was the only behaviour measured. For each phrase, the entry time was converted to entry speed using $(43 / 5) / (t / 60)$, where 43 is the size of the phrase in characters, 5 is the number of characters per word, t is the recorded entry time in seconds, and 60 is the number of seconds in a minute.[2]

2.5 Results/Discussion – Exp. 1

Counterbalancing the order of testing layouts achieved the desired outcome as the main effect and interactions for Group were not statistically significant. The grand mean for entry speed was 19.4 wpm. The speed for the Qwerty layout was quite fast at 26.5 wpm, while that for the Opti layout was only 12.3 wpm. The difference was statistically significant ($F_{1,10} = 797.0$, $p < .0001$).

There was considerable variation by participant. For Qwerty, participant means over the five trials ranged from 18.7 wpm to 30.2 wpm. For Opti, the means ranged from 6.7 wpm to 16.0 wpm. This suggests that participants approached the task with different attitudes on balancing speed with accuracy. The highest speeds recorded for single phrases were 35.0 wpm for Qwerty and 20.7 wpm for Opti.

There was also a significant effect for Trial ($F_{4,40} = 50.7$, $p < .0001$), implying that participants' entry speed increased with practice. The Layout by Trial interaction effect was also

---

[1] The following 45-character variant is sometimes used: "the quick brown fox jumped over the lazy dogs" [39]. In either case, the distinguishing feature is that the phrase contains every letter of the English alphabet. This ensures that every alphabetic key on the layout is tapped at least once. In fact, this phrase is somewhat atypical of English, since highly infrequent letters, such as "z", "x", and "q", are over-represented.

[2] It has been a convention since about 1905 to standardize the computation of entry speed in "words per minute", where a word is defined as five keystrokes [40, p. 182]. This includes letters, spaces, punctuation, and so on.

significant ($F_{4,40}$ = 2.7, p < .05), although much less so than either main effect. The trends by Layout and Trial are seen in Figure 2.

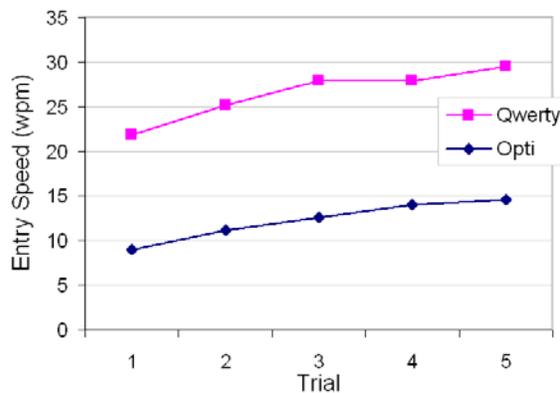Figure 2. Entry speed (wpm) by Layout and Trial for Experiment #1

3   Method – Experiment 2

The method for Experiment #2 was the same as for Experiment #1, except as follows.

Eighteen university-age students were involved as participants and assistants. None participated in the other experiments reported herein.

The comparison was between the Opti (Figure 1b) and Fitaly soft keyboard layouts. Fitaly is a product of Textware Solutions (textwaresolutions.com) and has a predicted expert entry speed of 42.0 wpm [24] (see Figure 3). None of the participants had previous experience with either keyboard layout.

Z	V	C	H	W	K
F	I	T	A	L	Y
space		N	E	space	
G	D	O	R	S	B
Q	J	U	M	P	X

Figure 3. Fitaly soft keyboard layout used in Experiment #2

Using Opti in both experiments serves as a calibration on the experimental methodology. The result for Opti should be similar in Experiments #1 and #2, as the apparatus and procedure were identical.

3.1   Results/Discussion – Exp. 2

As with Experiment #1, the Group main effect and interactions were not significant.

The layouts yielded similar results: 11.7 wpm for Opti and 12.2 wpm for Fitaly. The difference was not statistically significant ($F_{1,16}$ = 1.739, p > .05). As with Experiment #1, the main effect for Trial was significant ($F_{4,64}$ = 53.245, p < .0001), as was the Layout x Trial interaction ($F_{4,64}$ = 2.591, p < .05). The results are shown in Figure 4.
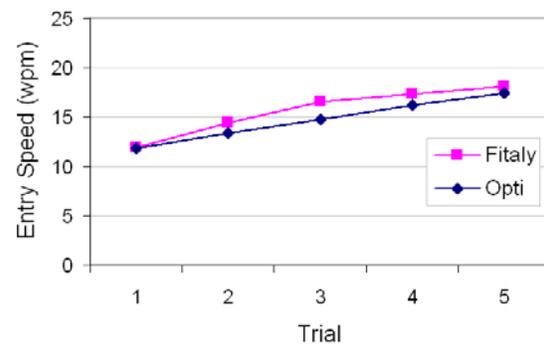

Figure 4. Entry speed (wpm) by Layout and Trial for Experiment #2

As seen in Figure 4, the significant Layout by Trial interaction appears as a jump in entry speed for the Fitaly layout in Trials 4 and 5. An explanation of this lies in comments made by a few participants after the experiment. Most participants felt they were faster with the Fitaly layout, even though the difference was not statistically significant overall. The reason cited by participants was that the words "jumps" and "dog" were very fast to enter due to key proximities (see Figure 5).
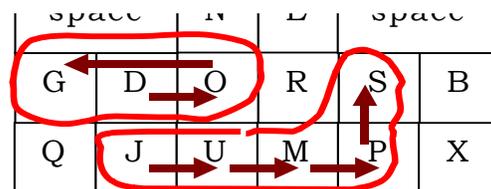

Figure 5. Fitaly input of "dog" (top) and "jumps" (bottom)

4 Method – Experiment 3

The method for Experiment #3 was the same as for Experiments #1 or #2, except as follows.

Twenty-four university-age students participated. None were involved in the other experiments reported herein. In this experiment, the comparison was between a Qwerty keyboard layout (Figure 1a) and a layout resembling a phone keypad (Figure 6).

1	2 ABC	3 DEF
4 GHI	5 JKL	6 MNO
7 PQRS	8 TUV	9 WXYZ
*	0 –	#

Figure 6. Phone keypad layout used in Experiment #3 and Experiment #4.

As a phone keypad is ambiguous for text entry, participants were given additional instructions to tap each key once only, and to assume automatic (and correct!) disambiguation is taking place.

4.1 Results/Discussion – Exp. 3

As with Experiment #1 and #2 the main effect and interactions for Group were not significant.

The entry speed was 48.3% faster for the Qwerty layout over the phone keypad layout, with mean of 28.1 wpm for Qwerty and 19.0 wpm for the phone keypad. As expected, the difference was statistically significant ($F_{1,22} = 65.80$, $p < .0001$).

The Trial main effect was significant ($F_{4,88} = 54.22$, $p < .0001$) indicating improvement with practice. This effect and the main effect for Layout are evident in Figure 7.
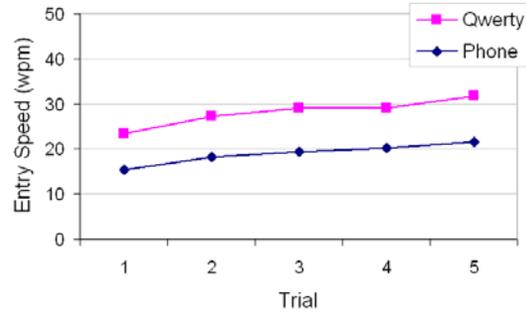

Figure 7. Entry speed (wpm) by Layout and Trial for Experiment #3

The Layout x Trial interaction was not significant ($F_{4,88} = 1.31$, $p > .05$) suggesting the improvement with practice was the same for both layouts.

5 Method – Experiment 4

The method for Experiment #4 was the same as for the previous experiments, except as follows.

Twenty-two university-age students were involved as participants and assistants. None participated in the other experiments reported herein.

In this experiment, the comparison was between a standard phone keypad (Figure 6) and a new layout called "QP" (Figure 8).

Figure 8. Qwerty-Phone (QP) layout used in Experiments #4 and #5.

The QP layout is a hybrid design. The three rows of letters on a Qwerty keyboard are mapped – without any re-ordering – to the three rows bearing letters on a phone keypad. The partitioning of letters was performed in a manner to minimize the Keystrokes Per Character (KSPC) statistic. For the design in Figure 8, KSPC = 1.0043, which makes the layout far less

ambiguous that the similarly-computed KSPC = 1.0072 for a phone keypad [22].

As in Experiment #3, since both phone keypads were ambiguous, participants were instructed to tap each key once only and assume disambiguation was working correctly.

5.1 Results/Discussion – Exp. 4

As with the previous experiments the main effect and interactions for Group were not significant. Contrary to expectations, the entry speed for the standard phone layout was massively (112%) faster than the QP layout, with a mean of 26.0 wpm for the standard phone against 12.3 wpm for the QP layout. The difference was statistically significant ($F_{1,22} = 48.95$, $p < .0001$).

The Trial main effect was significant ($F_{4,88} = 49.40$, $p < .0001$) indicating improvement with practice. This effect and the main effect for Layout are evident in Figure 12.
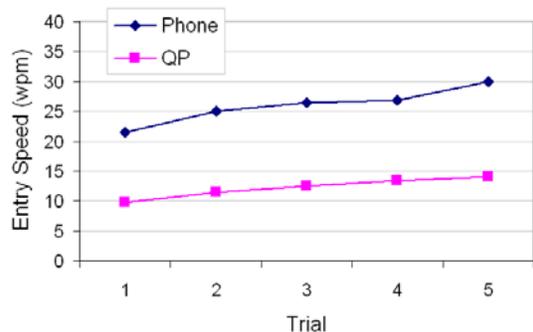

Figure 9. Entry speed (wpm) by Layout
and Trial for Experiment #4

The Layout by Trial interaction effect was also significant ($F_{4,88} = 4.50$, $p < .005$), suggesting that one method was easier to learn than the other.

6 Method – Experiment 5

The method for Experiment #5 was once again similar to the earlier experiments, except that in this case twelve university-age students were involved as participants (none had participated in the other experiments reported herein) and the comparison was between a Qwerty keyboard layout (Figure 1a) and the QP layout (Figure 8).

As with the preceding experiment, participants were instructed to tap each key once only, and to assume automatic disambiguation was taking place.

6.1 Results/Discussion – Exp. 5

Once again the main effect and interactions for Group were not significant.

The entry speed was 184% faster for the Qwerty keyboard layout over the QP layout, with mean of 34.5 wpm for Qwerty keyboard and 12.1 wpm for the QP layout. An analysis of variance revealed that the difference was statistically significant ($F_{1,9} = 80.78$, $p < .0001$).

The Trial main effect was significant ($F_{4,36} = 30.10$, $p < .0001$) indicating improvement with practice. This effect and the main effect for Layout are evident in Figure 10.
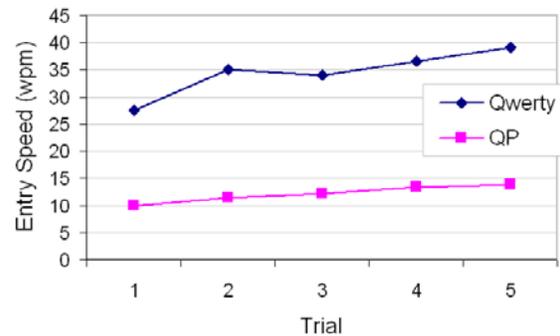

Figure 10. Entry speed (wpm) by Layout
and Trial for Experiment #5

The Layout x Trial interaction was also significant ($F_{4,36} = 6.97$, $p < 0.0001$) suggesting that improvement with practice was different for both layouts.

7 Summary and Discussion

In this section we provide an overall critique of the experimental methods by examining and comparing the results in the five experiments, and by comparing the results with other published research.

7.1 Combined Results

The results for the layouts tested in each of experiment are combined in Figure 11.
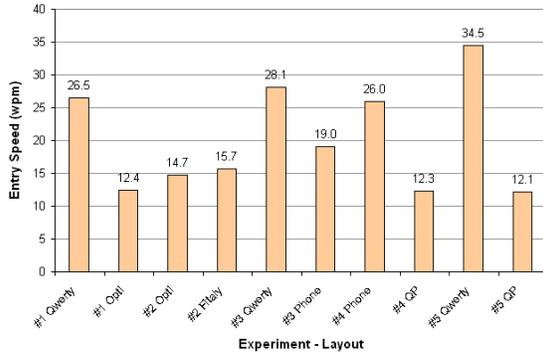
Figure 11. Entry speed (wpm) for each
layout for all five experiments

There are four important points of comparison: the three Qwerty keyboard tests (Experiments #1, #3, and #5), the two Opti tests (Experiment #1 and #2), the two Phone tests (Experiment #3 and #4), and the two Qwerty-phone (QP) tests (Experiment #4 and #5). For Qwerty, the means were 34.5, 28.1, and 26.5 wpm. These differences, especially the larger wpm reported in Experiment #5, may be due to prior experience. The subjects in Experiment #5 were all reported high daily usage of computers. The phone keypad differences could be similarly explained, as the group in Experiment #5 reported higher levels of text input than those in Experiment #4.

In general, the "between study differences" reported here are not surprising, especially considering the "between participant differences" which, for example, ranged from 18.7 wpm to 30.2 wpm for Qwerty in Experiment #1 and from 19.1 to 45.6 wpm for Qwerty in Experiment #5. Similar variations are reported in other studies. As just one example of this, phone keypad entry rates have been reported from as low as 7.9 words per minute [15] to as high as 21.0 wpm [36]. Variation across studies with the more novel keyboards, Opti and the QP layout, were much less. This reflects the effect of the participants having had no experience with the layout.

7.2   Are the Comparisons Valid?

This research is motivated to test a simple method for evaluating soft keyboard layouts. While the goal of simplicity is clearly met, the method is only useful if the results bear scrutiny in comparison with those using a more realistic apparatus and a more thorough procedure.

Fortunately, such a comparison is possible. MacKenzie and Zhang [26] compared the Qwerty and Opti soft keyboard layouts in a longitudinal experiment using custom experimental software and a Wacom tablet and stylus. The phrases of text presented to participants to enter were selected randomly from a set of 70 phrases. The average phrase was 25 characters in length.
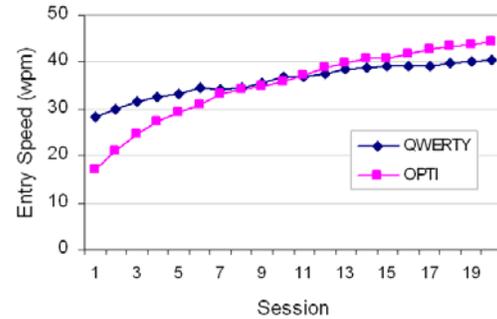

Figure 12. Qwerty vs. Opti results for entry
speed from MacKenzie and Zhang [26]

The comparison of relevance here is with the session one of MacKenzie and Zhang's results. The important statistics are summarized in Table 1.

Table 1.
Comparison With Session One Results
from MacKenzie and Zhang [26]

Study	Phrases	Testing Time (minutes)	Layout	
			Qwerty	Opti
MacKenzie & Zhang	50-60	20-22	28	17
Current	5	3-5	26.5	12.3

While the results for the Qwerty layout are quite close between the two studies, the results for Opti are on the low side: 12.3 wpm in Experiment #1 compared with 17 wpm in MacKenzie and Zhang's study. For both layouts, further improvement with practice seems likely, given the trends in Figure 2. So, higher figures seem reasonable for Experiment #1, had testing continued for 20-22 minutes, as in MacKenzie and Zhang's study. Since the rate of improvement in the current study was greater with Opti (for reasons noted earlier), the mean would likely be proportionally higher than for the Qwerty layout, perhaps settling in at the 17 wpm figure reported by MacKenzie and Zhang.

It seems reasonable to conclude, therefore, that the results observed in the current study are consistent with those reported by MacKenzie and Zhang. In fact, if the goal is to measure walk-up entry rates, limiting the test to 3-5 minutes of input is arguably preferable. Results so gathered are more representative of "walk-up" use than those obtained over 20-22 minutes of testing.

7.3 Critiquing the Method

While the empirical results are reasonable (see above), they are limited since only one dependent variable was used. The method is clearly a compromise, and is not presented here as a substitute for a full and proper empirical study.

On the plus side, the experiment design and implementation are straight-forward, since a physical computing device is not used, nor is any software written. Furthermore, the procedure takes only about 40 minutes, since participants are gathered and tested together. There is a bit of confusion and noise while the mockup sheets are distributed and the procedure explained, but this can be corrected with careful advance planning. It would be useful, for example, for the experimenter to enlist an assistant to distribute the mockup sheets. Having the participants serve as assistants in gathering measurements seemed to work quite well. Once the experiment was underway, participants proceeded without distractions.

It is extremely important that the experimenter create the correct atmosphere for the experiment. The present experiments seemed to succeed, as students took to the experiment seriously. In addition, the experiments stimulated debate about keyboard layouts as well as experiment design.

7.4 Measuring Accuracy

While accuracy was not measured in this experiment, it may be possible to modify the procedure to capture errors. If participants use a felt-tip pen, their taps will leave a mark on the paper. Following the input of a phrase, the paper could be inspected for errors. The process would be tedious, and perhaps error prone in itself. Additionally, such a method requires a fresh keyboard rendering for each trial. However, this modification is something to consider for future use of the method. One potential benefit is that

participants inclined to proceed recklessly, might be more careful if a record of their performance is generated.

8 Conclusions

We have demonstrated the use of paper mockups and hand timing to test soft keyboard layouts. Participants were simultaneously tested, and also served as assistants in conducting the experiment. We measured 26.5, 28.1 and 34.5 wpm for a Qwerty layout, 12.3 and 14.7 wpm for an Opti layout, 15.7 wpm for the Fitaly layout, 12.1 and 12.3 wpm for the QP phone layout, and 19.0 and 23.0 wpm for a phone keypad layout. Where comparisons could be made with more formal evaluations, results are reasonably consistent, suggesting that the methodology is useful as a quick and efficient means to empirically test soft keyboard layouts.

9 Acknowledgement

References

1. Aliakseyeu, D. and Martens, J.-B., The electronic paper prototype with visual interaction enriched windows, *Proceedings of the 2nd European Union symposium on Ambient Intelligence*, (New York: ACM, 2004), 11-14.

2. Andersson, G., Hook, K., Mourao, D., Paiva, A., and Costa, M., Using a Wizard of Oz study to inform the design of SenToy, *Proceedings of the Conference on Designing Interactive Systems - DIS '04*, (New York: ACM, 2002), 349-355.

3. Chandler, C. D., Lo, G., and Sinha, A. K., Multimodal theater: Extending low fidelity paper prototyping to multimodal applications, *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems - CHI '02*, (New York: ACM, 2002), 874-875.

4. Dahlback, N., Jonsson, A., and Ahrenberg, L., Wizard of Oz studies: Why and how, *Proceedings of the 1st International*

Conference on Intelligent User Interfaces - IUI '93, (New York: ACM, 1993), 193-200.

5. Goldstein, M., Bretan, I., Sallnas, E.-L., and Bjork, H., Navigational abilities in audial voice-controlled dialogue structures, *Behaviour & Information Technology*, *18*, 1999, 83-95.

6. Gong, J. and Tarasewich, P., Alphabetically constrained keypad designs for text entry on mobile phones, *Proceedings of the ACM Conference on Human Factors in Computing Systems - CHI 2005*, (New York: ACM, 2005), 211-220.

7. Gopher, D. and Raij, D., Typing with a two-hand chord keyboard: Will the QWERTY become obsolete?, *IEEE Transactions of Systems, Man, and Cybernetics*, *18*, 1988, 601-609.

8. Grady, H. M., Web site design: A case study in usability testing using paper prototypes, *Proceedings of the 18th Annual ACM Conference on Computer Documentation*, (New York: ACM, 2000), 39-45.

9. Hanington, B. M., Interface in form: Paper and product prototyping for feedback and fun, in *Interactions*, vol. 13: New York: ACM, 2006, January, 28-30.

10. Hendry, D. G., Mackenzie, S., Kurth, A., Spielberg, F., and Larkin, J., Evaluating paper prototypes on the street, *Extended Abstracts of the ACM Conference in Human Factors in Computing Systems - CHI '05*, (New York: ACM, 2005), 1447-1450.

11. Hoysniemi, J., Hamalainen, P., and Turkki, L., Wizard of Oz prototyping of computer vision based action games for children, *Proceeding of the Conference on Interaction Design and Children - IDC 2004*, (New York: ACM, 2004), 27-34.

12. Hughes, D., Warren, J., and Buyukkokten, O., Empirical bi-action tables: A tool for the evaluation and optimization of text input systems: Application I: Soft keyboards, *Human-Computer Interaction*, *17*, 2002, 271-309.

13. Hunter, M., Zhai, S., and Smith, B. A., Physics-based graphical keyboard design, *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems -*

CHI 2000, (New York: ACM, 2000), 157-158.

14. Hwang, S. and Lee, G., Qwerty-like 3x4 keypad layouts for mobile phone, *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems - CHI 2005*, (New York: ACM, 2005), 1479-1482.

15. James, C. L. and Reischel, K. M., Text input for mobile devices: Comparing model prediction to actual performance, *Proceedings of the ACM Conference on Human Factors in Computing Systems - CHI '01*, (ACM Press, 2001), 365-371.

16. Klemmer, S. R., Sinha, A. K., Chen, J., Landay, J. A., Aboobaker, N., and Wing, A., Suede: A wizard of oz prototyping tool for speech user interfaces, *Proceedings of the ACM Symposium on User Interface Software and Technology -- UIST 2000*, (New York: ACM, 2000), 1-10.

17. Kroemer, K. H. E., Operation of ternary chorded keys, *International Journal of Human-Computer Interaction*, *5*, 1993, 267-288.

18. Lewis, J. R., Kennedy, P. J., and LaLomia, M. J., Development of a digram-based typing key layout for single-finger/stylus input, *Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting*, (Santa Monica, CA: HFES, 1999), 415-419.

19. Lewis, J. R., LaLomia, M. J., and Kennedy, P. J., Evaluation of typing key layouts for stylus input, *Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting*, (Santa Monica, CA: HFES, 1999), 420-424.

20. Lewis, J. R., Potosnak, K. M., and Magyar, R. L., Keys and keyboards, in *Handbook on human-computer interaction*, (M. Helander, T. K. Landauer, and V. Prabhu, Eds.). Amsterdam: Elsevier, 1997.

21. Liu, L. and Khooshabeh, P., Paper or interactive?: A study of prototyping techniques for ubiquitous computing environments, *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems - CHI '03*, (New York: ACM, 2003), 1030-1031.

22. MacKenzie, I. S., KSPC (keystrokes per character) as a characteristic of text entry

techniques, *Proceedings of the Fourth International Symposium on Human-Computer Interaction with Mobile Devices*, (Heidelberg, Germany: Springer-Verlag, 2002), 195-210.

23. MacKenzie, I. S., Nonnecke, R. B., Riddersma, S., McQueen, C., and Meltz, M., Alphanumeric entry on pen-based computers, *International Journal of Human-Computer Studies*, *41*, 1994, 775-792.

24. MacKenzie, I. S. and Soukoreff, R. W., Text entry for mobile computing: Models and methods, theory and practice, *Human-Computer Interaction*, *17*, 2002, 147-198.

25. MacKenzie, I. S. and Zhang, S. X., The immediate usability of Graffiti, *Proceedings of Graphics Interface '97*, (Toronto: Canadian Information Processing Society, 1997), 120-137.

26. MacKenzie, I. S. and Zhang, S. X., The design and evaluation of a high-performance soft keyboard, *Proceedings of the ACM Conference on Human Factors in Computing Systems - CHI '99*, (New York: ACM, 1999), 25-31.

27. MacKenzie, I. S. and Zhang, S. X., An empirical investigation of the novice experience with soft keyboards, *Behaviour & Information Technology*, *20*, 2001, 411-418.

28. MacKenzie, I. S., Zhang, S. X., and Soukoreff, R. W., Text entry using soft keyboards, *Behaviour & Information Technology*, *18*, 1999, 235-244.

29. MacKenzie, I. S., Zhang, X. I., and Soukoreff, R. W., Stylus tapping on a soft keyboard, *Behaviour & Information Technology*, *18*, 1998, 235-244.

30. Maulsby, D., Greenberg, S., and Mander, R., Prototyping an intelligent agent through Wizard of Oz, *Proceedings of the ACM Conference on Human Factors in Computing Systems - CHI '93*, (New York: ACM, 1993), 277-284.

31. Molin, L., Wizard-of-Oz prototyping for co-operative interaction design of graphical user interfaces, *Proceedings of the Third Nordic Conference on Human-Computer Interaction - NordiCHI 2004*, (New York: ACM, 2004), 425-428.

32. Nielson, J., Paper versus computer implementations as mockup scenarios for heuristic evaluation, *Proceedings of IFIP INTERACT '90: Human-Computer Interaction*, (Berlin: Springer, 1990), 315-320.

33. Norman, D. A. and Fisher, E., Why alphabetic keyboards are not easy to use: Keyboard layout doesn't much matter, *Human Factors*, *24*, 1982, 509-519.

34. Ryu, H. and Cruz, K., LetterEase: Improving text entry on a handheld device via letter reassignment, *Proceedings of the 19th Conference of the Computer-Human Interaction Special Interaction Group (CHISIG) of Australia*, (New York: ACM, 2005), 1-10.

35. Sefelin, R., Tscheligi, M., and Giller, V., Paper prototyping - what is it good for?: A comparison of paper- and computer-based low-fidelity prototyping, *Extended Abstract of the ACM Conference on Human Factors in Computing Systems - CHI '03*, (New York: ACM, 2003), 778-779.

36. Silfverberg, M., MacKenzie, I. S., and Korhonen, P., Predicting text entry speed on mobile phones, *Proceedings of the ACM Conference on Human Factors in Computing Systems - CHI 2000*, (New York: ACM, 2000), 9-16.

37. Slaughter, I., Oard, D. W., Warnick, V. I., Harding, J. L., and Wilderson, G. J., A graphical interface for speed-based retrieval, *Proceedings of the 3rd ACM International Conference on Digital Libraries -- DL '98*, (New York: ACM, 1998), 305-306.

38. Snyder, C., *Paper prototyping: The fast and easy way to design and refine user interfaces*: San Francisco: Morgan Kaufmann, 2003.

39. Soukoreff, W. and MacKenzie, I. S., Theoretical upper and lower bounds on typing speeds using a stylus and soft keyboard, *Behaviour & Information Technology*, *14*, 1995, 370-379.

40. Yamada, H., A historical study of typewriters and typing methods: From the position of planning Japanese parallels, *Journal of Information Processing*, *2*, 1980, 175-202.

41. Zhai, S., Hunter, M., and Smith, B. A., The Metropolis keyboard: An exploration of

quantitative techniques for graphical keyboard design, *Proceedings of the ACM Symposium on User Interface Software and Technology - UIST 2000*, (New York: ACM, 2000), 119-128.

42. Zhai, S., Hunter, M., and Smith, B. A., Performance optimization of virtual keyboards, *Human-Computer Interaction*, *17*, 2002, 229-269.

43. Zhai, S., Sue, A., and Accot, J., Movement mode, hits distribution and learning in virtual keyboarding, *Proceedings of the ACM Conference on Human Factors in Computing Systems - CHI 2002*, (New York: ACM, 2002), 17-24.