An empirical investigation of the novice experience with soft keyboards

I. SCOTT MACKENZIE† and SHAWN X. ZHANG‡

†Dept. of Computer Science, York University, Toronto, Ontario, Canada M3J 1P3;
e-mail: smackenzie@acm.org

‡Corel Corporation, 1600 Carling Avenue, Ottawa, Ontario, Canada K1Z 8R7; e-mail: shawnz@acm.org

Abstract. An experiment with 12 participants tested text entry rates on two sizes of soft keyboards with either a Qwerty layout or a layout presenting a randomized letter arrangement after each tap. The randomized layout simulated the novice experience by requiring users to visually scan the layout for each tap to find the intended letter. Rates for the Qwerty layouts were about 20 wpm with no significant difference between the large and small size. Rates for both sizes of the randomized layouts were very low, about 5.4 wpm. This is the expected walk-up text entry rate with a soft keyboard bearing an unfamiliar layout. This empirical result allows us to reject a previous model of novice interaction that used Fitts' law for stylus movement and the Hick–Hyman law for visual scan time.

1. Introduction

Developing fast and efficient means for text entry is one of the most pressing research topics in today's frantic race for new mobile communications products. Despite the obvious appeal of miniaturization and, hence, portability, mobility bears a price. The physical means for input are significantly constrained by the lack of a keyboard and mouse—the traditional input devices for desktop systems. Other mechanisms must be explored, such as speech input, physical keyboards with fewer or smaller keys, or stylus input.

Using the cell phone market as an example, the growth of text messaging is nothing short of remarkable. In December 2000, for example, 15 billion short messaging service (SMS) messages were sent (see www.gsmworld.com). Most of these are phone-to-phone, although desktop-to-phone messaging is also supported. SMS messaging includes several appealing features. For one, SMS messages are very inexpensive compared to a voice call. In addition, the service is 'store and forward'. Hence, communication is asynchronous: The recipient may be offline when a message is sent. Finally, SMS messaging is unobtrusive. In situations where voice calls are inappropriate (e.g. during a group meeting), it is often possible to discretely read, and potentially respond to, an incoming SMS message.

As remarkable as the volume of SMS messages, is the woefully inadequate means to input them. Overwhelmingly, SMS messages are input using the telephone keypad and the multi-tap input technique (Detweiler *et al*. 1990, Silfverberg *et al*. 2000), although other techniques are emerging such as Tegic's *T9* (www.tegic.com), or Eatoni's *WordWise* (www.eatoni.com). Evidently, the appeal of text messaging is so strong that users are willing to 'put up with' the poor affordances of the 12-key phone keypad just to get their messages through.

Efficient text input is viewed as important, not only for cell phones, but also for a myriad of other mobile products likely to emerge in the near future. A common form factor is the personal digital assistant (PDA), such as the *PalmPilot*, which uses a stylus as the primary input device. Mobile information services such as web browsing, e-mail and online chat all require text input, and so the problem of supporting this form of interaction looms large. Most PDA text input is performed either through handwriting with automatic recognition or through an on-screen graphical representation of a keyboard, often called a 'soft keyboard'. This paper focuses on the latter.

1.1. *Skill transfer from touch typing*

Although the Qwerty layout is entrenched for physical keyboards, soft keyboards are easy to implement and customize. Therefore, the arguments for using a Qwerty layout are not as strong for soft keyboards, and a few alternative designs have emerged, such as *Fitaly* (www.textwaresolutions.com), *OPTI* (MacKenzie and Zhang 1999) and *Metropolis* (Zhai *et al.* 2000).

For soft keyboards, the arguments supporting the Qwerty layout are further diminished because touch-typing skill will not necessarily transfer to 'touch tapping'. The motor skill of two-handed, eyes-free touch typing is very different from the simple act of one-handed, eyes-on tapping with a stylus on a soft keyboard. That is, if we compare two touch typists— one 25 wpm and one 75 wpm—in their ability to use a soft keyboard, it is not certain that the 75 wpm typist would be faster, or substantially faster, than the 25 wpm typist.

Furthermore, it may emerge that fast touch typists are worse off than slow hunt-and-peck typists when the task is stylus tapping on a soft keyboard. The reason is this: touch typing is a highly learned (viz. automatic) motor act, whereas slow typing, such as hunt-and-peck, tends to be more visually guided. Slow typists may fair better because their style of conscious interaction keeps them thinking about the arrangement of letters on the keys.

Either way, the extent of skill transfer from touch typing to stylus tapping is an open question. We explore this question here in two contexts: tapping on a soft keyboard with a new layout and tapping on a soft keyboard with a Qwerty layout.

1.2. *The novice experience*

Most work on the design of soft keyboards focuses on the potential, or expert, text entry rate of a particular design. In this paper, we explore the opposite end of the learning curve: the novice experience. We are motivated by the belief that the novice experience is paramount for the success of new text input methods. This is at least partially due to the target market. Mobile products, such as cell phones and PDAs—once specialized tools for professionals—are increasingly targeted for the consumer market. It follows that 'immediate usability' is important. In other words, it may be a moot point to establish the expert, or 'potential', text entry rate for an input technique, if prolonged practice is required to achieve it. Consumers may be 'turned off' by their initial experience and frustration. The research literature on stylus-based text entry is replete with novel techniques that have not advanced into commercial products, at

least in part, due to the significant time required to attain proficiency in use (e.g. Goldberg and Richardson 1993, Mankoff and Abowd 1998, Perlin 1998, Venolia and Neiberg 1994).

As it turns out, measuring immediate usability is easier said than done. In typical studies of new interaction techniques, participants are given a demonstration of the technique followed by a brief practice session. Then, data collection proceeds over several blocks of trials, typically lasting 30–60 minutes. Measurements so gathered are by no means representative of expert behaviour. This requires prolonged testing! However, the measurements are also a poor indicator of novice behaviour—at least, in the sense of immediate, or walk-up, usability. Within a few minutes, participants' knowledge of the interaction technique develops and the novice status fades. At the very least, performance measures so gathered are not indicative of walk-up use if measured after 30–60 minutes of practice.

An earlier study in mobile text entry sought to capture 'immediate' usability. MacKenzie and Zhang (1997) measured character-level accuracy with *Graffiti*, after tightly controlling users' initial exposure over 1-minute and 5-minute periods. Results indicated character-level accuracy of 86% within one minute and 97% within five minutes. Measurements were not gathered on text entry speed, however. The main point to note here is that new methodologies are emerging to evaluate this important aspect of interaction: walk-up, or immediate, usability.

1.3. *Modelling user interaction with soft keyboards*

Modelling user interaction with soft keyboards is easier than with physical keyboards. Although physical keyboards engage ten fingers, soft keyboards are generally used with a stylus. The motor component of soft keyboard interaction can be modelled using Fitts' law because the act of entering a character is akin to a target selection task. Entry rates for experts are predicted by computing the time to hit each key, given any previous key, and then applying a weighting factor based on probabilities in the language (see Soukoreff and MacKenzie 1995, MacKenzie *et al.* 1998, Lewis *et al.* 1999, MacKenzie and Zhang 1999, Hunter *et al.* 2000, Zhai *et al.* 2000).

However, interaction for novices is more complicated. The motor act of tapping a key with a stylus is trivial, and, in this sense, differences between novices and experts are probably small, perhaps negligible. However, knowledge of the layout—the letter arrangement on the keys—is a different matter. Users unfamiliar with the layout must visually scan the keyboard to find the

desired letter, and this takes time. How much time? Previous work (MacKenzie *et al*. 1998, Soukoreff and MacKenzie 1995) used the Hick–Hyman law (Hick 1952, Hyman 1953) for choice reaction time to estimate the time to locate one letter from a set of 27 (26 letters plus space):

$$RT = 0.2 \log_2(27) = 0.951 \text{ s} \qquad (1)$$

With this, novices were distinguished from experts simply by adding 951 ms to the tapping time for each key. For experts, the visual scan time was set at zero.

The following thought experiment further illustrates the impact of the visual scan time. If we assume visual scan time is the *only* component of the interaction—that the motor component is negligible—then the entry rate is

$$(1/0.951) \times (60/5) = 12.62 \text{ wpm} \qquad (2)$$

The expression above simply converts the visual scan time into its reciprocal, 'characters per second', and into words per minute by multiplying by 60 seconds/minute and dividing by five characters/word. So, 12.62 wpm represents an upper bound entry rate for novices confronting a new letter arrangement. The actual novice entry rate is lower by a factor representing the motor component of the interaction, which adds extra time for each character entered. Earlier work proposed a novice entry rate for a Qwerty layout of 8.9 wpm or 1.348 seconds/character. This figure is simply the sum of the visual scan time (0.951 s) and the mean stylus movement time (0.397 s) (Soukoreff and MacKenzie 1995). The stylus movement time was obtained using Fitts' law.

The novice entry rate and, in particular, the visual scan time metric have never been tested empirically for interaction with soft keyboards. One of the goals of this research is to provide such a test.

1.4. *Effect of keyboard size*

Although a few studies have examined the effect of keyboard size for *physical* keyboards (Drury and Hoffmann 1992), there has been little study for soft keyboards. An exception is the work of Sears and colleagues (Sears *et al*. 1993) who controlled soft keyboard size in a user evaluation. They found a text entry rate of 32 wpm for a large keyboard and 21 wpm for a small keyboard. However, the keyboard was rendered on a touch screen and users entered text using all fingers on both hands; so, the comparison with stylus tapping on a soft keyboard is weak.

We feel keyboard size is worthy of investigation because of an apparent gap between intuition and theory. On the one hand, small soft keyboards should be harder to use than large soft keyboards because the targets are smaller. Alternatively, perhaps small soft keyboards will be easier to use because less pen travel is require. However, neither viewpoint is supported in theory, explained as follows. The task of tapping keys on a soft keyboard is a standard Fitts' law task. Yet, according to Fitts' law, keyboard size should not affect performance. This follows from the simple observation that scaling a keyboard up or down produces a 'net zero' effect on Fitts' index of difficulty:

$$ID = \log_2\left(\frac{A}{W} + 1\right) \qquad (3)$$

where A is the movement amplitude (distance) and W is the target width (size). The ratio A/W is insensitive to the size of the keyboard: Doubling the distance between keys also doubles the size of keys!

We feel it is important to empirically test these opposing views (intuition versus theory) for the limited case of stylus tapping on soft keyboards. So, keyboard size is included as a controlled variable in the experiment described in the following section.

Given the preceding, we undertook the design of an experiment. The goal was to provide an empirical understanding of following issues in soft keyboard design:

- Is the Hick–Hyman visual scan time prediction a useful metric in modeling novice behaviour with a soft keyboard?
- Does the skill of touch typing transfer to stylus tapping?
- Does the size of the keyboard affect the entry speed and accuracy of a soft keyboard?

2. Methodology

2.1. *Participants*

Twelve students and staff from a local university were paid to participate in the experiment. Six were male and six were female; all were right handed. Some were computer science students or staff, while others were in life science disciplines. All used computers on a daily basis. None were regular users of PDAs or used a stylus as an input device. They were recruited from those who responded to a posting to student and staff e-mail lists.

2.2. *Apparatus*

The experiment software was developed in C++ using Microsoft Windows for Pen Computing 1.1. A

Wacom PL-100V tablet was attached to the system. The Wacom tablet is both an LCD display and a digitizer. It has a 12-inch B&W backlit LCD display panel with 640×480 pixel resolution. Using the combination of the tablet and host computer enabled the experiment to run without system lag and allowed user entry to also appear on a regular VGA monitor. Instead of pressing and releasing a mouse button to 'click', users tap a stylus on the tablet surface to complete the same action. The stylus and tablet use the standard Windows mouse driver.

The software displayed two different sized keyboards for the text entry task. Each size also had two different layouts. One displayed the standard Qwerty arrangement. We call this the 'fixed' layout. The other had the same form factor but randomized the letter assignment to keys each time a key was tapped. We call this the 'random' layout. The random layout, we conjecture, maintains the novice status of the user because with each tap a new letter arrangement must be visually scanned. The keys on the large keyboard were 28×28 pixels each, or 10×10 mm measured on the display. The keys on the small keyboard were 18×18 pixels each, or 6.4×6.4 mm on the display.

The task implemented in software was to memorize and enter a short phrase of text. This approach simulates a text *creation* task in that the user knows exactly what to enter. This is in contrast to a text *copy* task wherein the user's focus of attention continually switches between the source text and the keyboard.

Figure 1a shows a session with the large fixed layout. The letter arrangement was fixed, as per the Qwerty standard. A sample phrase is displayed on the first line and user input is on the second line. Figure 1b shows a session with the large randomized keyboard layout. Note that letter assignment on the keys was randomized following each key tap. The experiment screen appeared the same for the two small keyboard conditions, except that each key was 10×10 pixels, as noted earlier.

2.3. *Design*

The experiment was a 2×2 within-subjects design. The factors and levels were as follows:

- Keyboard size {small, large}
- Keyboard layout {fixed, random}

There were four different combinations of layout and size. Each participant finished four blocks, one for each condition. Each block contained 10 phrases. The order of the conditions was balanced to reduce interactions between conditions.
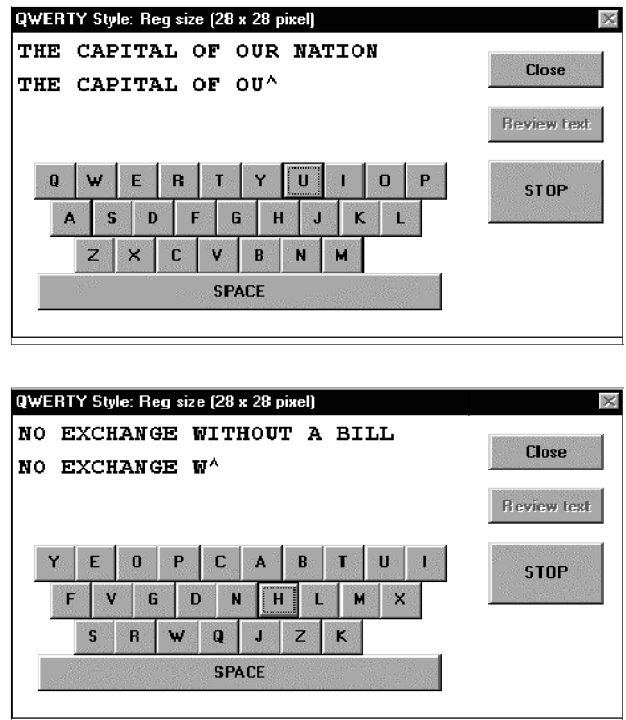

Figure 1. Experiment screen: (a) large fixed layout; (b) large randomized layout.

The text phrases were selected randomly from a sample set of 70 phrases (see Figure 1 for two examples). The phrases were crafted for ease of memorization. Phrases with infrequent letters (e.g. Q, X, Z) were also included to ensure the set was representative of common English. To test this, we compared the letter frequencies against those in a standard corpus (Maynzer and Tresselt 1965). The correlation was $r = .9516$.

2.4. *Data collection*

For each key tapped, the following information was collected:

- Given character
- Given character's position (ID on the keyboard)
- User entered character
- User entered character's position
- Time elapsed between characters (in ms)
- Error (0 = correct, 1 = error)

The first letter in each phrase was excluded from the timing measurement and from the calculation of the correlation of the letter frequencies in the test phrases with common English. This is because there is no

reference point from which to time the entry of the first character.

2.5. *Procedure*

The tablet was positioned on a separate desktop away from the experimenter's host machine. The room lights were dimmed to minimize the glare on the LCD panel. Prior to the formal test, all participants were given a typing test using *Typing Tutor IV* on an Apple *Macintosh*. The text for the typing test was modified from a longer story to fit on one screen and to show alphabetic characters with minimal punctuation.

After the typing test, participants were given written instructions explaining the task and device. They were specifically asked to focus on both speed and accuracy. They were instructed to ignore mistakes and to carry on with a phrase when a mistake was made. When an error occurred, a 'click' was heard through the system speaker.

They were then given the tablet and the stylus. The tablet was tilted off the desk to provide a good viewing angle (about 15°). It was also adjusted to have appropriate contrast and brightness. To provide a basic level of familiarity with the apparatus, participants were given one practice phrase (about 25 characters) prior to data collection.

In the formal test, one phrase was displayed at a time. With the fixed keyboard layout, the participants only needed to copy the text. The visual search time was presumed minimal since the layout was fixed as per a Qwerty keyboard, which all participants used on a daily basis. With the randomized layout, however, they had to visually search for the next letter on the keyboard after each letter was entered. The typing test was given again at the end of the last session.

3. Results and discussion

3.1. *Text entry speed*

This experiment measured text entry performance for 15–20 minutes of use only. Our focus was not on prolonged user learning and proficiency, but rather on users' initial experience with a soft keyboard—either with the well-known Qwerty layout or with an unfamiliar layout. The novice status of the participants was maintained for the random layouts, despite 15–20 minutes of usage, because a new letter arrangement appeared with each tap. Effects of keyboard size were also investigated.

The grand mean for entry speed was 13.00 wpm. Figure 2 shows the breakdown by condition. Clearly,
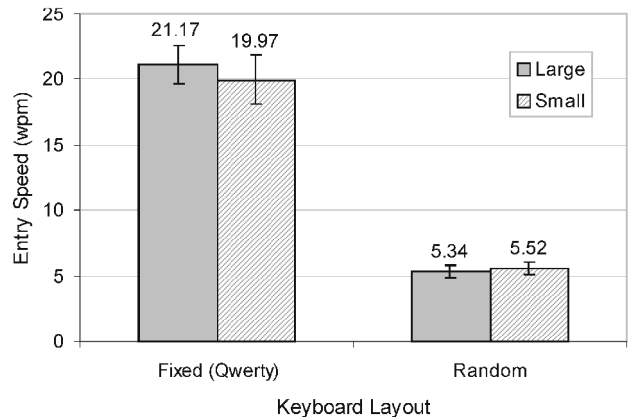

Figure 2. Entry speed (wpm) by keyboard size and layout (error bars show one standard deviation).

the main effect for layout was significant ($F_{1,11} = 314.5$, $p < .0001$). There was no effect for keyboard size ($F_{1,11} = 2.63$, $p > .05$); however, the interaction effect between keyboard layout and size was marginally significant ($F_{1,11} = 7.18$, $p < .05$).

The large fixed keyboard was slightly faster (21.17 wpm) than the small fixed keyboard (19.97 wpm). When the letter assignment was randomized after each tap, however, the smaller keyboard showed a slightly faster entry speed (5.52 wpm) than the larger one (5.34 wpm).

Finding no significant difference on entry speed for keyboard size is consistent with the Fitts' law prediction. That is, scaling a soft keyboard up or down, within limits, does not impact text entry speed.

The distance from participants' eyes to the tablet was about 30 cm, resulting in a field of about 12 degrees for the small keyboard and 17 degrees for the large keyboard. Although the difference is only 5 degrees, we observed with the large keyboard, that users' hands tended to block a relatively large area of the keyboard. Moving their hand away to regain the full view took longer as well. All these factors would contribute to the slower performance with the large, randomized layout. We will say more about this later.

Although participants were not instructed to anchor any body part, they tended to anchor their wrist on the rim of the tablet when using the fixed keyboard layout. With the randomized layout, they tended to anchor their elbow on the desk and hold their wrist in the air above the tablet.

3.2. *Error rates*

Error rates were quite low overall with a grand mean of 1.94% (see figure 3). There was a significant difference
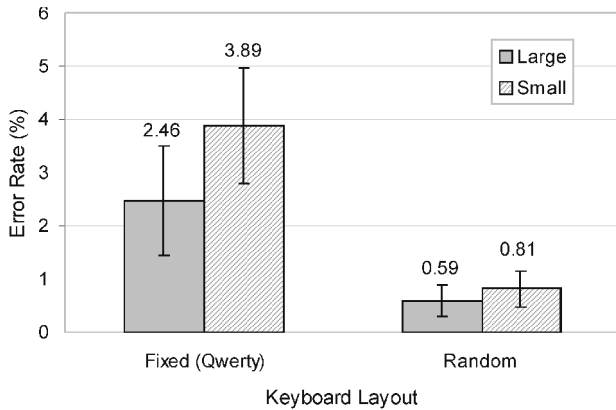
Figure 3. Error rate (%) by keyboard size and layout (error bars show one standard deviation).

in error rates between the two keyboard layouts ($F_{1,11} = 17.87$, $p < .005$). The fixed layouts had error rates of about 3.2% and the random layouts had error rates of only about 0.7%.

The low error rates for the random layouts are likely due to participants' constant need to visually acquire a letter before moving the stylus. This, we conjecture, resulted in a more cautious interaction style with the random layout than with the fixed layout. The confidence in being familiar with the fixed Qwerty layout allowed participants to proceed with quicker and more automatic motor control. This may be the source of the higher error rates for the fixed layouts.

The effect for keyboard size on error rate was also significant ($F_{1,11} = 3.42$, $p < .01$), as was the size by layout interaction ($F_{1,11} = 5.53$, $p < .05$). Finding a significant difference on error rates for keyboard size is contrary to the Fitts' law prediction, but is consistent with one intuitive viewpoint noted earlier, namely, that small soft keyboards should be harder to use than large soft keyboards, because the targets are smaller. However, the error rates were low for both sizes, so the lack of a significant difference in entry speed across the two sizes of keyboards is the more important finding.

3.3. *Visual scan time for interaction with a soft keyboard*

The mean entry rate of about 5.4 wpm for the soft keyboards with the randomized layouts is about 40% lower than the predicted novice rate of 8.9 wpm from another study (MacKenzie *et al*. 1998) cited earlier in this paper. A rate of 5.4 wpm corresponds to about $(1/5.4) \times (60/5) = 2.222$ seconds per character. If we extract the movement component predicted by Fitts' law, the result is $2.222 - 0.397 = 1.825$ seconds. This is about double the time expected of 0.951 s for the

visual scan time. Clearly, behaviour is present that is unaccounted for.

In fact, having now conducted an empirical test that maintains users' novice status (i.e. participants confront a new letter assignment with each tap), we are in a position to categorically reject the simple addition of the visual scan time metric to the Fitts' law stylus movement time as an adequate representation of novice behaviour with a soft keyboard. Note that our use of the term 'novice behaviour' simply implies that the user is working with a soft keyboard with an unfamiliar letter arrangement.

Our observations of participants' behaviour were as follows. While visually scanning the randomized keyboard layout, participants consistently raised the stylus after each tap. They held it mid-air, generally about 4–8 cm above the display. This behaviour seemed to occur for two reasons: first, to ensure the keys were not obstructed from view by the hand or stylus, and, second, to return the stylus to a neutral position from where the next movement would begin. Some participants simply swung the pen over the keyboard to follow the focal movement of eyes and tapped down as soon as the next letter was located. In either case, this is not a simple target selection task. This behaviour clearly weakens the potential of Fitts' law to model the movement component of the task for novices. Although the Hick–Hyman metric may still be valid in general, clearly as applied here it is confounded with the complex movement behaviour we observed.

3.4. *The novice-to-expert transition*

Our novice text entry rate of 5.4 wpm is a lower bound—a walk-up entry rate. In practice, a soft keyboard would have a fixed letter arrangement, and users would gain familiarity with it quickly. No study exists that specifically examines text entry rates for novices after controlled initial exposures of, say, one minute, five minutes, 15 minutes, etc.

In MacKenzie and Zhang's longitudinal study of a high performance soft keyboard, known as *OPTI*, entry rates were 17 wpm on the first session, rising to 45 wpm on the 20th session (MacKenzie and Zhang 1999). Each session lasted about 20 minutes and contained 5–11 blocks of text, each containing 10 phrases of about 25 characters. So, even the rate for the first session (17 wpm) is well past what we consider as a walk-up text entry rate. Their study had a different motivation, however, focusing on the expert potential of an optimized design.

The work presented here examined the opposite end of the learning spectrum. To garner support for new

designs of soft keyboards, however, we suggest that evaluations attend to both the expert potential and to the novice experience, the latter measured over controlled yet brief periods of walk-up exposure to the technology.

3.5. *Skill transfer from touch typing to stylus tapping*

Our participants' typing speeds on a physical keyboard averaged over the pre- and post-tests ranged from 19 to 74 wpm. Across participants, the soft keyboard entry rates varied from 15.4 to 25.1 wpm for the large fixed keyboard and from 13.1 to 26.0 wpm for the small fixed keyboard. The relationship by participant between typing speed and tapping speed for the four keyboards is shown in table 1.

There was a modest positive correlation between touch typing speed and stylus tapping speed on the soft keyboards with the fixed (Qwerty) layout: $r = 0.416$ for the large keyboard and $r = 0.523$ for the small keyboard. However, when entering text on soft keyboards with an unfamiliar layout, the correlation with typing speed was quite low: $r = 0.073$ and $r = 0.081$ for the large and small randomized layouts, respectively.

The simple interpretation is that, yes, on average a fast typist will tap faster on a soft keyboard than a slow typist, but only if the user is working with the familiar Qwerty layout. However, the correlations for the fixed layouts were not as high as commonly found, for example, in Fitts' law studies with pointing devices where correlations above 0.9 are common. Although familiarity does help, the two tasks are fundamentally different, both in the mental model and in the motor skills involved. This is support for our belief that more efficient soft keyboard design is warranted.

Figure 4 shows a scatter plot and trend line for the small fixed keyboard, which produced the highest correlation with touch typing speed. Note in the scatter plot that the slowest typist (19 wpm) also recorded the slowest stylus tapping rate (13.14 wpm) (P12 in table 1). On the other hand, P7 recorded the fastest typing rate (74 wpm) while achieving the second fastest stylus-tapping rate (23.96 wpm).

The relationship in the trend line is expressed by the following linear regression equation:

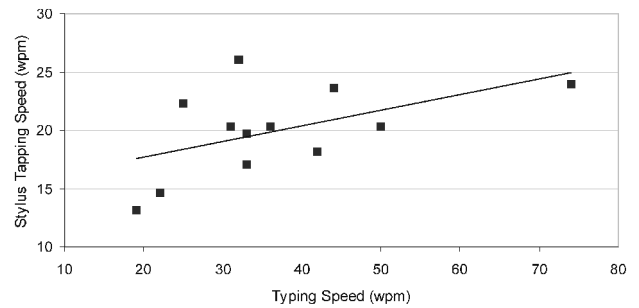$$Tapping\_Speed = 15.0 + 0.13\ Typing\_Speed \qquad (4)$$

Figure 4. Relationship between stylus tapping speed and typing speed for the small fixed (Qwerty) layout.

Table 1.   Entry rates (wpm) by participant and condition.

Participant	Typing	Tapping			
		Fixed/large	Fixed/small	Random/large	Random/small
P1	42	19.01	18.19	4.38	4.89
P2	44	21.74	23.61	5.71	6.35
P3	32	24.77	26.03	6.38	6.81
P4	50	20.33	20.28	3.80	4.12
P5	36	20.33	20.28	3.80	4.12
P6	33	21.25	17.09	6.34	6.57
P7	74	25.07	23.96	6.52	6.15
P8	22	18.98	14.68	5.19	5.34
P9	31	21.36	20.32	5.74	6.28
P10	33	20.25	19.70	5.26	5.50
P11	25	25.52	22.35	5.56	5.11
P12	19	15.44	13.14	5.35	4.95
Mean	36.8	21.17	19.97	5.33	5.52
SD	14.77	2.89	3.79	0.93	0.92
Correlation	−	0.416	0.523	0.073	0.081

Note: Correlations are with typing speed (2nd column).

Due to the small sample size, caution in using Equation 4 as a predictor is warranted.

There is no support for the earlier suggestion that slow typists—those tending to use visually-guided hunt-and-peck strategies—might fare better than fast typists when using a soft keyboard. Possible exceptions are P3 and P11. Their typing speeds were among the lowest (32 and 25 wpm, respectively), yet they produced among the highest stylus tapping speeds (26.03 and 22.35 wpm, respectively).

4. Conclusions

This experiment examined some important issues in soft keyboard usability. We tested two common sizes of soft keyboards and found no significant differences in entry speeds. We found a moderate positive correlation ($r \approx 0.5$) between touch-typing speed on a physical keyboard and stylus entry speed on a soft keyboard bearing a Qwerty layout. Thus, users experienced with desktop systems are expected to benefit from skill transfer when switching to a soft keyboard with a Qwerty layout. However, no skill transfer will take place ($r < 0.1$) if the layout is new or unfamiliar to users.

The use of a letter assignment that was randomized after each key tap allowed us to measure the walk-up text entry rates for soft keyboards with unfamiliar layouts. We found very low entry rates—about 5.4 wpm. This we offer as the expected walk-up entry rate for soft keyboards with unfamiliar layouts, such as recent designs claiming better-than-Qwerty rates for expert users.

Our results do not bode well for efforts to improve on the Qwerty layout by rearranging letters to minimize pen travel. Although such designs may yield a soft keyboard highly optimized for experts, the novice-to-expert transition cannot be ignored. In particular, the work presented here underscores the importance of measuring and acknowledging walk-up entry rates for new designs, and to measure and note these after specified intervals, such as one minute, 10 minutes or 15 minutes.

References

Detweiler, M. C., Schumacher, R. M. and Gattuso, N. L. 1990, Alphabetic input on a telephone keypad. In *Proceedings of the Human Factors Society 34th Annual Meeting* (Santa Monica, CA: Human Factors Society), pp. 212–216.
Drury, C. G. and Hoffmann, E. R. 1992, A model for movement time on data-entry keyboards. *Ergonomics*, **35**, 129–147.

Goldberg, D. and Richardson, C. 1993, Touch-typing with a stylus. In *Proceedings of the INTERCHI '93 Conference on Human Factors in Computing Systems*, 24–29 April 1993 (New York: ACM), pp. 80–87.
Hick, W. E. 1952, On the rate of gain of information. *Quarterly Journal of Experimental Psychology*, **4**, 11–36.
Hunter, M., Zhai, S. and Smith, B. A. 2000, Physics-based graphical keyboard design. In *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems – CHI 2000*, 1–6 April 2000 (New York: ACM), pp. 157–158.
Hyman, R. 1953, Stimulus information as a determinant of reaction time. *Journal of Experimental Psychology*, **45**, 188–196.
Lewis, J. R., Kennedy, P. J. and Lalomia, M. J. 1999, Development of a digram-based typing key layout for single-finger/stylus input In *Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting* (Santa Monica, CA: HFES), pp. 415–419.
MacKenzie, I. S. and Zhang, S. X. 1997, The immediate usability of Graffiti. In *Proceedings of Graphics Interface '97*, 21–23 May 1997 (Toronto: Canadian Information Processing Society), pp. 120–137.
MacKenzie, I. S. and Zhang, S. X. 1999, The design and evaluation of a high-performance soft keyboard. In *Proceedings of the ACM Conference on Human Factors in Computing Systems - CHI '99*, 15–20 May 1999 (New York: ACM), pp. 25–31.
MacKenzie, I. S., Zhang, X. I. and Soukoreff, W. 1998, Stylus tapping on a soft keyboard. *Behaviour & Information Technology*, **18**, 235–244.
Mankoff, J. and Abowd, G. A. 1998, Cirrin: a word-level unistroke keyboard for pen input In *Proceedings of the ACM Symposium on User Interface Software and Technology – UIST 98*, 1–4 November 1998 (New York: ACM), pp. 213–214.
Maynzer, M. S. and Tresselt, M. E. 1965, Table of single-letter and digram frequency counts for various word-length and letter-position combinations. *Psychonomic Monograph Supplements*, **1**, 13–32.
Perlin, K. 1998, Quikwriting: continuous stylus-based text entry. In *Proceedings of the ACM Symposium on User Interface Software and Technology – UIST '98* (New York: ACM), pp. 215–216.
Sears, A., Revis, D., Swatski, J., Crittenden, R. and Shneiderman, B. 1993, Investigating touchscreen typing: the effect of keyboard size on typing speed. *Behaviour & Information Technology*, **12**, 17–22.
Silfverberg, M., MacKenzie, I. S. and Korhonen, P. 2000, Predicting text entry speed on mobile phones. In *Proceedings of the ACM Conference on Human Factors in Computing Systems – CHI 2000*, 1–6 April 2000 (New York: ACM), pp. 9–16.
Soukoreff, W. and MacKenzie, I. S. 1995, Theoretical upper and lower bounds on typing speeds using a stylus and soft keyboard. *Behaviour & Information Technology*, **14**, 370–379.
Venolia, D. and Neiberg, F. 1994, T-Cube: a fast, self-disclosing pen-based alphabet. In *Proceedings of CHI'94 Conference on Human Factors in Computing Systems*, 24–28 April 1994 (New York: ACM), pp. 265–270.
Zhai, S., Hunter, M. and Smith, B. A. 2000, The Metropolis keyboard: an exploration of quantitative techniques for graphical keyboard design. In *Proceedings of the ACM Symposium on User Interface Software and Technology - UIST 2000*, 5–8 November 2000 (New York: ACM), pp. 119–128.