Text Entry Using Five to Seven Physical Keys

Elliot Lockerman, Shuobi Wu
Carnegie Mellon University
Pittsburgh, PA, USA
{elockerm, shuobiw}@andrew.cmu.edu

Ariel Rao
Facebook, Inc.
Menlo Park, CA, USA
arielrao@fb.com

Jarret Lin, Neil Bantoc, Brad A. Myers
Carnegie Mellon University
Pittsburgh, PA, USA
{jjlin, nbantoc}@andrew.cmu.edu;
bam@cs.cmu.edu

***Abstract*—We designed, implemented, and evaluated a small physical keyboard, composed of four to five keys with a variety of mappings for the letters, along with one or two navigational function keys operated by the thumb. The small size allows it to be used for smartwatch text entry, and can be mounted on various body parts as a wearable keyboard. It is primarily used as an ambiguous keyboard, akin to T9, although a multi-tap mode is also present. Our keyboard layouts include Alphabetic, Collapsed QWERTY, Mnemonic (where letters are grouped based on their shapes), and Optimized (which minimizes the number of conflicts). In a between-users study, participants achieved an average of 15.4 words per minute (WPM) across all layouts, with one user with the Collapsed QWERTY layout reaching a top speed of 30.6 WPM after 4 hours of practice. User feedback was generally favorable.**

Keywords— Text entry; mobile input; one-handed keyboard; interaction techniques; handheld device; smart watch.

I. Introduction

Text entry techniques on ultra-small devices has thus far been limited. Most input methods use soft keyboards with small buttons, leading to the "fat finger" problem, in which a finger covers multiple buttons, with an increase in input errors [1]. Furthermore, the lack of tactile feedback limits the top speed [2].

We introduce WatchBoard (Figure 1), a compact physical wearable keyboard with a small number of keys supporting a variety of layouts, intended primarily for smartwatches or wearable devices. The keyboard has 4 or 5 letter keys, depending on the layout, plus one or two keys operated by the thumb which are used for space, backspace and selection. Furthermore, depending on the mode, the desired letter for a key press can be indicated manually via multi-tap or predicted by an algorithm, similar to T9 [3]. The current implementation is designed to be worn on the left arm/hand for typing with the right hand. In a commercial design, the buttons might slide underneath the watch or be mounted on the watch band.

In summary, the contributions of this work include:

- A novel hardware design for text entry using a small number of buttons attached to a smartwatch or other wearable display.
- Four mappings of English letters to the keys which enable multi-tap or recognizer-based text entry. All of the mappings include novel aspects.
- A user study of the physical design and four layouts which showed that people can learn to use this text entry mechanism. The C-QWERTY layout proved fastest, with a top speed of 30.6 words per minute (WPM). These rates are comparable or better than reported for other small devices [3] [4].

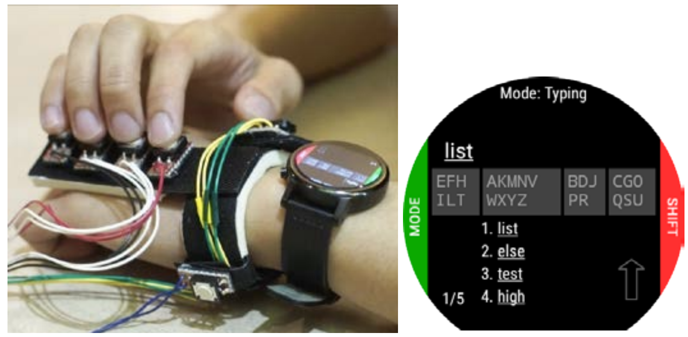

Figure 1: WatchBoard Prototype in Use, and close up of screen

II. Related Work

A. Phone Input Techniques

Traditional feature phones map letters alphabetically to the 12 keys on the keypad and allow two different input techniques: multi-tap and T9 [3]. In the multi-tap mode, repeated taps in quick succession cycle through the letters of a key. In T9, a single tap is made on the key mapped to the target letter and statistical models are used predict the target word from sequences of key presses. These typically show around 8 to 15 WPM for multi-tap and around 9 to 20 WPM for T9 [5] [4] [6].

B. Smartwatch Input Techniques

Komnios and Dunlop [7] present a soft keyboard for smartwatches containing six ambiguous keys, each representing three to six letters, laid out in alphabetical order. The open source tool OpenAdaptxt was used to predict words given a series of button presses. Swiping gestures were used for handling special characters and cycling through predictions. In a series of user studies, participants averaged 8.1 WPM. They concluded that keyboards with a small number of ambiguous keys represented a viable approach to smartwatch text entry. However, using soft keyboards has the disadvantage that a high level of eye-hand coordination is required to type in the absence of tactile feedback, and utilizing multiple fingers is difficult, additionally limiting the top speed.

WatchWriter [8] represents the current state of the art for smartwatch text entry. Similar to a typical smartphone soft keyboards, it uses a QWERTY layout, allowing for both taps and gesture typing. Despite its small size, participants achieved an average of 22 (tapping) and 24 wpm (gestures).

978-1-5386-0443-4/17/$31.00 ©2017 IEEE

C. Non-traditional Keyboard Layouts

Keyboards based on letter-shape mnemonics include UniWatch [9], a 3-key keyboard designed for smartwatches based on UniGlyph [9], a 4-button key-based text entry technique. In these two keyboards, each letter is grouped based on its shape (curve, stroke, or loop). Letters can be entered by pressing one of three buttons, by drawing the shapes, or with a swipe at an angle. The word is disambiguated by a statistical algorithm. Both techniques were evaluated using a French sentences set, with unknown amount of practice. UniWatch achieved 13.78 WPM at maximum, while UniGlyph achieved 17.36 WPM at maximum.

Another approach is to begin with a QWERTY keyboard and "compress" it. Half-QWERTY [10] is a one-handed keyboard, consisting of one half of a QWERTY keyboard. When pressing and holding the spacebar, the keyboard is mirrored, so the other half can be typed. In a user study, an average speed of 35 wpm was reached after 10 sessions, with high variability. However, at the last session, all subjects were still improving, and the error rate was double that of two handed typing.

1line [11] and Minuum's mini mode [12] use a vertically compressed QWERTY keyboard, i.e., each column (such as Q, A, and Z), is represented by a single button. Which letter was desired is determined by statistical analysis, with a menu providing alternatives. While 1line was developed for iPad, allowing use of all fingers, Minuum was designed for phones and thumb typing. After five 20 minute typing session, 1line reported an average typing speed of just over 30 wpm. No evaluation is available for Minuum.

A layout can also be optimized by some physical, psychological, or linguistic metric. The OPTI layout [13] is a soft keyboard for mobile text entry designed with typing behavior in mind. Keys are also distributed across the keyboard based on how frequently they are used and the other characters they co-occur with in English. Under simulation, typing with OPTI was predicted to be 35% faster than typing with QWERTY, at an upper bound entry rate of 58.2 wpm.

Other smartwatch text entry techniques report speeds around 9-16 WPM [14] [15]. A huge variety of other systems have been developed for a plethora of circumstances, with highly varying results, e.g., [16] [17].

III. WATCHBOARD DESIGN

A. Hardware

The buttons consisted of four or five letter keys as well as two function keys: space/select, and backspace. Menu selection was accomplished using the letter keys (described in Usage, below). Each button is independently mounted on the harness with Velcro allowing for individual adjustment of the spacing and positions.

B. Software

A watch testing application was developed for the Moto 360 (2nd Generation) smartwatch. It displayed the current mode, the target sentence, layout guide for the layout in use, and the list of potential matches (see Figure 1). In addition to typing on the physical keyboard, the user could swipe up and down to scroll through pages of potential matches, swipe left to capitalize the next letter (indicated on the screen with a shift arrow), and swipe right to reveal the mode pane.

To bootstrap the statistics software's core corpus, we used Nolls' top 3000 American English words [18]. We then used the Enron Mobile Email dataset to train our model, since the style and formality of mobile email has reasonable overlap with messaging text.

The prediction software takes in a key mapping as well as a sequence of key presses *S* and returns a list of candidate words in order of maximum likelihood. To determine the ranking of this list, all words that *S* prefixes are retrieved. Each candidate word is then scored as a unigram as well as a bigram and trigram with previously entered words. A length penalty is also applied to prefer shorter words.

C. Usage

Three modes are available for all keyboard layouts: Typing mode, Manual mode, and Symbol mode. Typing Mode is semantically similar to T9, requiring the user to press a single letter key for each desired letter, which produces a list of predicted words to select from. The list of words is updated every time the user presses a new key. To select a word, first the select button is pressed. Then, the letter key in the same ordinal position as the desired option is pressed. Pressing backspace instead of a letter key cancels the selection process. Manual mode and Symbol mode allow for entering arbitrary letters and symbols, respectively, using multitap, akin to a feature phone.

D. Layouts

The *Alphabetic* layout (see Figure 2a) was designed based on alphabetic order. We hypothesized that since everyone is familiar with alphabetic order, this might have a faster initial learning, but possibly a slower top speed. The Alphabetic layout has 4 letter keys, one for each of the non-thumb fingers. Much like the 12-key phone keypad, we shrunk the number of keys by grouping letters in adjacent order. Boundaries were elected to reduce the number of collisions between letters that are difficult to distinguish between (e.g., "b" and "f" in "bind" and "find").

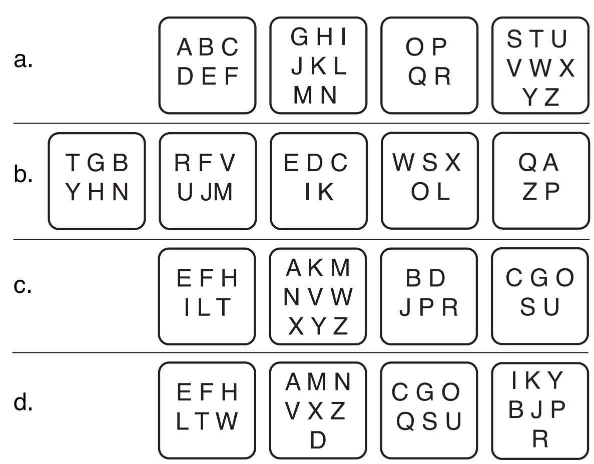

Figure 2: Layouts. a. Alphabetic, b. C-QWERTY, c. Mnemonic, d. Statistically Optimized. In addition, all layouts have two function keys (space/select and backspace)

The *Collapsed-QWERTY* (aka *C-QWERTY*) layout (Figure 2b) is spiritually a combination of the approaches taken by the Half-QWERTY [10] and 1line [11]/Minuum [12] keyboards: the keyboard is both compressed vertically and mirrored horizontally. The main invariant we preserved was the finger-to-letter mapping from traditional QWERTY keyboards. We hypothesized that people who touch type using the traditional ten finger method on a QWERTY keyboard would have fast learning with this layout, but only after an adjustment period.

The *Mnemonic* layout (Figure 2c) was originally inspired by the Wubizixing input method (aka Wangma method) [19], in which strokes of Chinese letters are grouped into different letters on a QWERTY keyboard. We used similar shape-based heuristics, so that letters are grouped by the shape of their capitals into four groups (thus requiring only 4 buttons): (1) only horizontal and vertical lines, (2) only lines but at any angle (3) a horizontal or vertical line followed by a curve and (4) mainly curves. We hypothesized that shape-based categorization would be easy to remember, leading to faster initial learning.

The *Optimized* layout (Figure 2d) was statistically designed to reduce collisions—i.e., given a series of button presses, the number of possible resulting words is minimized. This was accomplished by greedy character selection which considers letter frequency, word patterns, and finger strength. We hypothesized a long initial learning time due to unfamiliarity with the layout, but higher maximum speeds due to reduced collisions.

IV. Evaluation

A. Study 1

We first performed a between-subjects study examining the initial learning period for all of the layouts. This served as a pilot to help narrow down which layouts were worthy of a longer longitudinal study.

1) Participants

13 participants (ages 21-19, avg 24; 1 female) were recruited for study 1, and 12 of them completed an entire session. One participant was dismissed before completing the study, and his data was discarded. He was arching his wrist while typing, and complained of hand pain. His typing position did not improve after coaching, and the pain persisted, leading to termination of the session. Each participant was randomly assigned to a layout for the duration of the study. All of the participants were recruited from our university, and were paid $20 for their time.

2) Study Design

The session started with a one-minute typing test on a traditional QWERTY laptop keyboard, where the subjects had an average of 42 WPM, ranging from 29 to 62. This was followed by one practice in Manual mode and one in Typing mode, and then 12 trials in Typing mode. Each trial consisted of 3 minutes of text entry, and we measured how many words they were able to enter correctly during that time. Short breaks were allowed between each trial; longer breaks were allowed every three trials. During trials seven through nine, the layout guide was hidden to encourage memorization; participants were allowed to "peek" at the guide by tapping the watch screen. The participants were not tested or instructed on the use of shift, paging through the options list, or changing modes. A handful of words were not typeable in Typing mode due to the desired word not appearing on the first page of options shown on the watch, in which case the participants were instructed to skip such words when they occurred. (Such words could be entered through scrolling the word list or manual entry, but these were known to be much slower.)

TABLE 1. Summary Statics for Study 1, Regular Trials

Layout	Mean	Std. Dev	Min	Max
Alphabetic	8.3	2.1	5.2	12.8
C-QWERTY	7.9	2.0	4.9	12.5
Mnemonic	7.3	1.5	4.4	10.3
Optimized	7.1	1.9	3.1	10.7

All trials consisted of typing a series of phrases from the Enron Mobile Email phrase set [20]. The phrases were first striped of punctuation, made lowercase, and combined so there were at least 100 words available. All participants were given the same phrases in the same order. Subjects were encouraged to correct errors (colored read) as they typed, and the results include the time to correct errors (e.g., using the backspace key or gesture).

3) Results

The 12 participants produced a total of 106 test trials. On average, 7.68 WPM was reached across all layouts (SD = 1.92 WPM). The averages for the different layouts were 8.34 WPM on Alphabetic layout, 7.94 WPM on C-QWERTY, 7.34 WPM on Mnemonic layout, and 7.08 WPM on Optimized layout (see Table 1). With $\alpha = 0.05$, a one-way analysis of variance (one-way ANOVA) on WPM yielded no significant difference among keyboard layouts, $F(3,102) = 2.463$, $p = 0.067$.

B. Study 2

The layouts which seemed the fastest (Alphabetic and C-QWERTY) were chosen for study 2. We ran a between-subjects longitudinal study, where each participant was assigned to one layout, and had 4 one-hour sessions separated by 24 to 48 hours since using a separation between 1 and 2 days has been shown to support learning over the course of a week [21].

1) Participants

16 right-handed participants (ages 18-31, avg 24; 6 female) were recruited for our study (an additional one was not able to return and their data was not used). The average traditional QWERTY laptop keyboard speed was 81 WPM (min: 51, max: 133). Each participant was randomly assigned a layout for the duration of the study. All of the participants were recruited from local universities, and were given $80 for their time for all 4 sessions, paid at the end of the last session.

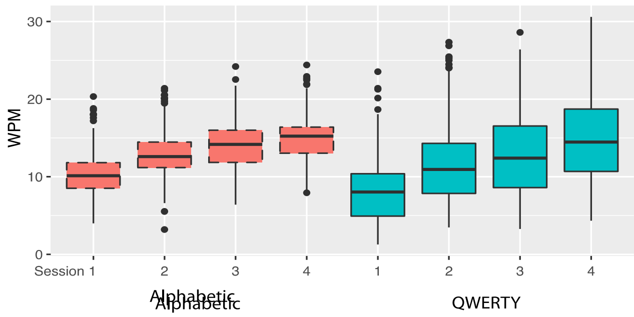

Figure 3: WPM by Keyboard Layout and Session.

2) Study Design

The study protocol was largely the same as study 1's, except that study 2 was longitudinal, with 4 sessions. Additionally, the hidden trials were moved to be at the end of the sessions instead of in the middle to lessen interference with the visible sessions.

3) Results

The 16 participants produced a total of 765 test trials, with 576 visible trials, and 189 hidden. The data were normal for Alphabetic and nearly normal for C-QWERTY. The two groups have almost equal sample size (n-alphabetic=383, n-C-QWERTY=382), and unequal variances (V-alphabetic= 14.79263, V-C-QWERTY=38.56446). So Welch's t-tests were used to analyze whether there are significant differences between group means. The trend for each layout across the four sessions is shown in Figure 3. In visible trials, Alphabetic layout achieved an average of 13.34 WPM, while C-QWERTY layout achieved an average 12.51 WPM. With $\alpha = 0.05$, a Welch two Sample T-test on WPM yielded very little significant difference between the two keyboard layouts, $p = F_{473.2}(2.0195) = 0.044$ across all visible trials in all sessions. Participants generally corrected errors as they typed (so the reported speeds include the correction times), and the uncorrected error rate was very low (a few percent).

A post hoc Tukey test showed all pairwise comparisons between any two sessions with each layout was significant. For the C-QWERTY layout, the difference between 1st and 2nd session is significant, with $p = 0.007$. The difference between the first and last session is very significant, with $p < 10^{-7}$. For Alphabetic layout, the difference between 1st and 2nd session is significant, with $p < 10^{-6}$. The difference between the first and last session is very significant, with $p < 10^{-7}$. The learning curves (Figure 6) do not appear to have converged after 4 sessions, so we expect that with additional practice, speed will continue to improve.

In addition, we were interested in the data from the last session because participants had had a reasonable amount of practice by then. Looking at the last session alone (Table 2), participants achieved 14.80 WPM for all visible trials (SD = 4.47 WPM). The averages for the different layouts were 14.29 WPM on Alphabetic layout, and 15.23 WPM on C-QWERTY. With $\alpha = 0.05$, a Welch two Sample T-test on WPM yielded no significant difference between the two keyboard layouts on the last session, $p = F_{78.671}(-1.215) = 0.228$. In hidden trials, participants reached an average of 14.39 WPM (SD = 5.57 WPM).

TABLE 2. SUMMARY STATISTICS FOR STUDY 2 VISIBLE TRIALS, LAST SESSION ONLY

Layout	Mean	StDev	Min	Max
Alphabetic	14.1	2.0	9.1	17.3
C-QWERTY	15.2	5.8	8.4	28.6

A two-way ANOVA on average speed on the last session vs. speed of typing on a traditional keyboard (levels: above median and below median) found a significant interaction. In particular, a Tukey's HSD post-hoc found that for both layouts, participants who have lower than median laptop keyboard typing speed are significantly slower than those who are above median, with $p < 10^{-7}$. The difference between the two layouts in terms participants who are below median is also significant, $p < 0.001$. However, the difference between two layouts in terms of participants who are above median is not significant, $p = 0.06$.

V. DISCUSSION

On average, participants for Alphabetic and QWERTY achieved similar top speeds (Table 2). However, there was more variability for Collapsed QWERTY, and the maximum attained speed was far higher for C-QWERTY – over 30 WPM vs 24.4 WPM for Alphabetic (although this was not during the last session). This may indicate that C-QWERTY is superior for certain groups, i.e., strong typists on regular QWERTY. In contrast, for Alphabetic, the increase was far lower, suggesting that the gains are due to procedural memories of the QWERTY layout, and not factors generic to keyboard use. This validates the efficacy of skill transfer with regards to the C-QWERTY keyboard.

VI. LIMITATIONS AND FUTURE WORK

There are several limitations which motivate future work. First, we were only able to collect data over four sessions. It is likely that most subjects did not attain their asymptotic fastest speed by the conclusion of the study. We therefore believe that with further practice, higher speeds would be achieved.

Second, the C-QWERTY does not benefit all users equally: it seemed to work best for good touch typists, but those who hunt-and-peck (or use "incorrect" fingers) did not perform as well. Future work would include investigating other layouts and better training protocols.

VII. CONCLUSIONS

We have presented a new way to do text entry for smart watches or other wearable devices, using a small number of keys on a physical keyboard, to be used by one hand. Additionally, we have demonstrated the utility of skill transfer from traditional keyboards to one of our layouts. As more small devices require user input, our research provides data points about wearable text entry mechanisms that might be considered.

REFERENCES

[1] S. S. Arefin Shimon, C. Lutton, Z. Xu, S. Morrison-Smith, C. Boucher, and J. Ruiz, "Exploring Non-touchscreen Gestures for Smartwatches," *Proc. 2016 CHI Conf. Hum. Factors Comput. Syst.*, pp. 3822–3833, 2016.

[2] E. Hoggan, S. A. Brewster, and J. Johnston, "Investigating the Effectiveness of Tactile Feedback for Mobile Touchscreens," *Proc. SIGCHI Conf. Hum. Factors Comput. Syst. (CHI '08)*, pp. 1573–1582, 2008.

[3] C. James and K. Reischel, "Text input for mobile devices: comparing model prediction to actual performance," *Proc. SIGCHI Conf. Hum. Factors Comput. Syst. (CHI '01)*, pp. 365–371, 2001.

[4] I. S. MacKenzie, H. Kober, D. Smith, T. Jones, and E. Skepner, "LetterWise: prefix-based disambiguation for mobile text input," *Proc. 14th Annu. ACM Symp. User interface Softw. Technol.*, vol. 3, no. 2, pp. 111–120, 2001.

[5] J. Wang, S. Zhai, and J. Canny, "SHRIMP: solving collision and out of vocabulary problems in mobile predictive input with motion gesture," *Proc. 28th Int. Conf. Hum. factors Comput. Syst. (CHI '10)*, pp. 15–24, 2010.

[6] J. O. Wobbrock, D. H. Chau, B. a Myers, and M. G. Hall, "An Alternative to Push, Press, and Tap-tap-tap: Gesturing on an Isometric Joystick for Mobile Phone Text Entry," *Proc. Int. Conf. Hum. factors Comput. Syst. (CHI '07)*, pp. 667–676, 2007.

[7] A. Komninos and M. Dunlop, "Using a smart-watch as a input device for text," *IEEE Pervasive Comput.*, vol. 13, no. 4, 2014.

[8] M. Gordon, T. Ouyang, and S. Zhai, "WatchWriter: Tap and Gesture Typing on a Smartwatch Miniature Keyboard with Statistical Decoding." *Proc. Int. Conf. Hum. factors Comput. Syst. (CHI '16)*, pp. 3817-3821.

[9] F. Poirier and M. Belatar, "Uniwatch - some approaches derived from uniglyph to allow text input on tiny devices such as connected watches," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2015, pp. 554–562.

[10] E. Matias, I. S. Mackenzie, and W. Buxton, "Half-QWERTY: Typing With One Hand Using Your Two-handed Skills," *Conf. Companion Hum. Factors Comput. Syst. (CHI'94)*, pp. 51–52, 1994.

[11] F. C. Y. Li, R. T. Guy, K. Yatani, and K. N. Truong, "The 1line keyboard: A QWERTY layout in a single line," *Proc. 24th Annu. ACM Symp. User interface Softw. Technol. - UIST '11*, pp. 461–470, 2011.

[12] Whirlscape Inc., "Minuum." [Online]. Available: http://minuum.com/.

[13] S. I. Mackenzie, S. X. S. X. S. X. Zhang, I. S. MacKenzie, and S. X. S. X. S. X. Zhang, "The Design and Evaluation of a High-Performance Soft Keyboard," *Proc. SIGCHI Conf. Hum. Factors Comput. Syst. (CHI'99)*, May, pp. 25–31, 1999.

[14] S. Oney, C. Harrison, A. Ogan, and J. Wiese, "ZoomBoard: a diminutive qwerty soft keyboard using iterative zooming for ultra-small devices," *Proc. CHI 2013*, pp. 2799–2802, 2013.

[15] J. Hong, S. Heo, P. Isokoski, and G. Lee, "SplitBoard : A Simple Split Soft Keyboard for Wristwatch-sized Touch Screens," *Proc. CHI '15*, pp. 1233–1236, 2015.

[16] X. A. Chen, T. Grossman, and G. Fitzmaurice, "Swipeboard: A Text Entry Technique for Ultra-Small Interfaces That Supports Novice to Expert Transitions," *Proc. UIST '14*, pp. 615–620, 2014.

[17] P.-O. Kristensson and S. Zhai, "SHARK^2: A Large Vocabulary Shorthand Writing System for Pen-Based Computers," *Proc. 17th Annu. ACM Symp. User interface Softw. Technol. - UIST '04*, vol. 6, no. 2, pp. 43–52, 2004.

[18] P. and B. Noll, "Nolls' Top 3,000 American English Words." [Online]. Available: https://web.archive.org/web/20050309021251/http://www.paulnoll.com/China/Teach/English-3000-common-words.html.

[19] "Wangma." [Online]. Available: http://www.wangma.net.cn/.

[20] K. Vertanen and P. Kristensson, "A versatile dataset for text entry evaluations based on genuine mobile emails," *MobileHCI '11 Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, pp. 295–298, 2011.

[21] N. J. Cepeda, N. Coburn, D. Rohrer, J. T. Wixted, M. C. Mozer, and H. Pashler, "Optimizing distributed practice theoretical analysis and practical implications," *Exp. Psychol.*, vol. 56, no. 4, pp. 236–246, 2009.