Few-key Text Entry Revisited: Mnemonic Gestures on Four Keys

Jacob O. Wobbrock, Brad A. Myers and Brandon Rothrock
Human-Computer Interaction Institute
Carnegie Mellon University
Pittsburgh, PA 15213 USA
{ jrock, bam, rothrock }@cs.cmu.edu

ABSTRACT
We present a new 4-key text entry method that, unlike most few-key methods, is gestural instead of selection-based. Importantly, its gestures mimic the writing of Roman letters for high learnability. We compare this new 4-key method to predominant 3-key and 5-key methods theoretically using KSPC and empirically using a longitudinal study of 5 subjects over 10 sessions. The study includes an evaluation of the 4-key method without any on-screen visualization—an impossible condition for the selection-based methods. Our results show that the new 4-key method is quickly learned, becoming faster than the 3-key and 5-key methods after just ~10 minutes of writing, although it produces more errors. Interestingly, removing a visualization of the gestures being made causes no detriment to the 4-key method, which is an advantage for eyes-free text entry.

ACM Classification Keywords
H.5.2. Information interfaces and presentation: User interfaces — *Input devices and strategies.*

Author Keywords
Text entry, text input, date stamp, selection keyboard, gestures, unistrokes, mobile devices, wearables, EdgeWrite.

INTRODUCTION
Text entry with only a few keys has been studied in mobile computing for some time [1,6,7]. As mobile technologies shrink while becoming more powerful and network-aware, 3-to-5 key methods remain relevant for devices such as wrist-watch PDAs, GPS units, and 2-way pagers. Areas besides mobile computing may also benefit from few-key methods. For example, fabric keypads have been sewn into smart clothing [9], and few-key methods have been used for people with limited ranges of motion [2].

Prior few-key methods have mostly been based on selection, displaying letters on a screen (Figure 1). Although selection methods are easy to learn, they have serious drawbacks for mobile text entry: (1) they require a screen; (2) the selections themselves consume precious screen real-estate; (3) they require a user's visual attention and cannot be done by feel; (4) they involve two foci-of-attention, the text being written and the selectable letters; and (5) they can be quite slow and tedious.

CHI 2006, April 22-27, 2006, Montréal, Québec, Canada.

Figure 1. A classic 3-key design. The letter "e" is currently selected.

Gestural methods, on the other hand, depend not on an on-screen depiction of selections but on the execution of meaningful motor patterns. Therefore, they generally do not suffer from the aforementioned drawbacks. One drawback of gestural methods, however, is that gestures must be learned. Thus, learnability is crucial [12]. In the few cases where gestures have been applied to keys [2,3,4], they have been somewhat arbitrary and difficult to learn. However, the advantages of gestures for few-key text entry warrant the investigation of a quickly learnable gestural technique. We present a new 4-key method that relies on mnemonic gestures reminiscent of Roman letters. Our method adopts the EdgeWrite alphabet [13], originally designed for use on PDAs by people with tremor.

PRIOR FEW-KEY TEXT ENTRY METHODS
3-Key Date Stamp
Text entry with three keys has been studied in-depth [6,7]. Three-key methods rely on two keys to move a selector left and right and a third key to select a letter. Although there are many possible layouts, we chose one that a previous study found "particularly promising" [7] (Figure 1). This design places space to the left, and the selector snaps there after each entry. This enables users to unhesitatingly move the selector to the right after each entry. It has a keystrokes per character (KSPC)[1] of 10.53 using the frequencies from [11]. Layouts with lower KSPC exist, but these often increase visual search time at the expense of speed [1,7]. The design in Figure 1 was measured at 9.10 words per minute (WPM) and 2.11% errors in a single session [7]. Note that we augmented this design with backspace ('<') for error correction. Its placement to the left of space does not increase KSPC because the selector does not wrap.

[1] Keystrokes per character (KSPC) is a frequency-weighted measure of how many key-presses are required to enter each letter [6].

The 3-key method is greatly enhanced by key-repeat. Our key-repeat times were taken from prior work [7] and set to 176 ms for the initial delay and 32.1 ms for the repeat delay. These are fast key-repeat times and enable high performance by experts.

5-Key Selection Keyboard

The 5-key method uses four keys to move over a matrix of letters and a fifth key for selection [1,6]. Our design is based on the Glenayre *AccessLink II* pager [6] (Figure 2). Its alphabetic layout is optimized to put common letters 'e', 's', 't', 'n', 'o' and 'u' near space. Where the actual device had punctuation marks we replaced them with asterisks since punctuation was not used in our other methods.

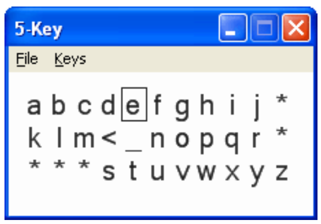

Figure 2. The 5-key method used in our study. This alphabetic layout is based on the Glenayre *AccessLink II* two-way pager.

The KSPC for the commercial product is 3.13. As with the 3-key method, we augmented the 5-key method with backspace ('<'). Its placement to the left of space minimally increases KSPC to 3.24. Like the 3-key method, the 5-key method employs snap-to-space ('_') and key-repeat.

Both the 3- and 5-key designs could reduce KSPC using optimization techniques. But KSPC is not the sole design factor, as visual search time is significant [1]. Manufacturers have generally eschewed fully optimized keyboards, perhaps favoring users' first impressions over their extended performance. Certainly, the intermittent use of few-key text entry methods requires that they be quickly learnable. Our new 4-key method also prizes initial usability, employing mnemonic Roman-like gestures over short arbitrary strokes that would be faster to perform.

Gestural Few-Key Methods

There have only been a few gestural few-key text entry methods. One method, *MDITIM* [3], is not intended specifically for keys but is adaptable to them. Its letters are comprised of strokes in the four cardinal directions and are generally not similar to Roman forms. We calculated its KSPC to be 3.06. Its inventors report that entry with the four keyboard arrow keys was ~4.9 WPM and ~3% errors after ten 30-minute sessions (~5 hours) using a stylus version [3].

Another technique, *UDLR* [2], is intended for the four keyboard arrow keys. All letters consist of 3 arrow-key-presses, i.e., KSPC is 3.00. After a week of practice, speeds reached ~13.5 WPM, but users reportedly suffered from high confusion due to the arbitrary key sequences.

Finally, Jannotti's design for *Iconic* text entry [4] uses all 10 keys of a numeric keypad, allowing users to trace letters that are somewhat reminiscent of Roman forms. *Iconic*'s KSPC is 2.43. It was reported to have a Fitt's law-derived theoretical speed of 19.8 WPM, but to our knowledge, this was never empirically validated.

4-KEY EDGEWRITE

EdgeWrite is a unistroke method originally designed for stylus entry on PDAs for people with tremor [13]. With a stylus, strokes are performed inside a square hole with stabilizing plastic corners and edges. EdgeWrite gestures are fully defined by the sequence of corner-hits they make inside the square. In our new 4-key version, each key-press represents a corner-hit (Figure 3). Thus, unlike MDITIM or UDLR, key-presses do not represent strokes, but endpoints of strokes, allowing for more Roman-like gestures.

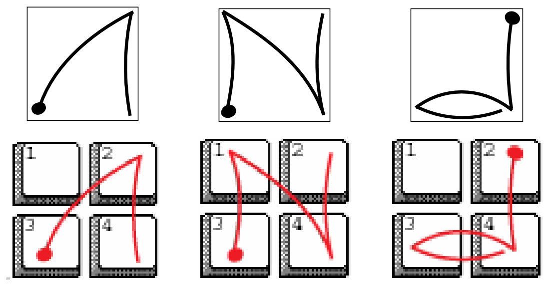

Figure 3. EdgeWrite letters mapped to four keys. Letters are defined by their sequence of corner-hits: i.e., "a" = 324, "n" = 3142, "d" = 2434. For a full character chart, see www.edgewrite.com.

The KSPC of EdgeWrite's primary letter forms is 3.52.[2] Admittedly, this is higher than the 5-key selection keyboard (3.24). But gestural methods do not require visual search like selection methods do. Thus, in this case, we have a tradeoff between KSPC and visual search.

Adapting a unistroke stylus method to four keys requires a solution to the segmentation problem, since "stylus lift" is not relevant. For segmentation we chose an adaptive timeout that adjusts on a per-letter basis to the speed at which a user makes a letter according to Equation 1.

$$T = F \cdot \frac{\sum_{i=1}^{n-1} t_{i+1} - t_i}{n-1} \qquad (1)$$

Here, T is the time until segmentation, F is a multiplier preference ranging from 1.20 (expert) to 2.00 (novice), t_i is the time of the i^{th} key-press, and n is the number of key-presses ($n > 1$). Thus, for users making fast strokes, segmentation occurs sooner than for users making slow strokes. After each key-down event, the timer is stopped so that segmentation cannot occur until all keys are up. When all keys are up, the timeout is computed and the timer is restarted. When the timer elapses, segmentation occurs, meaning the corner sequence is recognized and reset.

[2] EdgeWrite defines alternate forms for most letters to increase guessability [12]. Its primary forms are those considered most intuitive. Using its shortest forms, KSPC drops minimally to 3.30.

EVALUATION

Our study compared our new 4-key technique with the prior 3-key and 5-key methods. We conducted the study over 10 short sessions simulating "daily intermittent use."

Method

Five subjects (2 female) ranging from 27 to 33 years old took part in our study over 10 consecutive days. Subjects were all right-handed daily computer users. Their mean keyboard typing rate was 70.79 WPM (σ=13.76) with 3.54% total errors (σ=1.22%). None of the subjects had used any of the few-key methods. They were paid $10 per session.

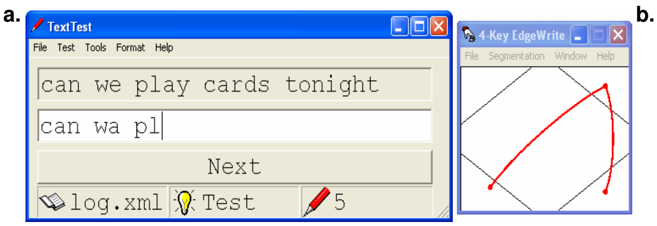

Figure 4. (a) The text entry test software. **(b)** 4-key EdgeWrite. During testing, methods were placed below the phrases.

The test software (Figure 4a) presented phrases from [8]. In all, it logged 1600 test phrases, or about 50,000 characters. Backspace was supported, and subjects were not forced to remain synchronized with the presented text [10]. As in prior work [1], the numeric keypad was used with one hand to control the few-key methods. Subjects were told to use the key configuration they found most comfortable. They used 3 fingers with 3-key, 3 or 4 fingers with 5-key, and 4 fingers with 4-key (e.g., on the 1, 2, 4 and 5 keys). Using the numeric keypad allowed us to compare methods under ideal conditions: one-handed with familiar comfortably-sized keys. Actual devices could be tailored toward any of the techniques should results warrant further design.

Each session consisted of 2 warm-up phrases and 8 test phrases for each of the four methods. These 10 phrases took ~5 minutes to complete per method. To see how the gestural method fared with no visualization, we included a condition *4-key-noviz*, where the stroke window (Figure 4b) was removed. Although this resulted in two EdgeWrite methods per session, the total time for EdgeWrite was still just ~10 minutes. Moreover, the two selection methods shared common features, e.g., key-repeat times and certain keys. The presentation order of the four methods was counterbalanced according to a Latin Square.

During the 2 warm-up phrases, an EdgeWrite character chart was displayed, but this chart was *not* shown during the 8 test phases. Not surprisingly, subjects had to guess many strokes in the early sessions [12]. Although not showing a character chart undoubtedly hurt the 4-key methods, we wanted to assess learnability without aids.

Results

We analyzed our repeated measures data using a random-effects model in which *Session* (1..10) and *Method* (3-key, 5-key, 4-key, 4-key-noviz) were repeated factors. We included *Method Order* (1..4) but found no order effects, indicating adequate counterbalancing. Measures of a single subject were not independent, so *Subject* was modeled as a random effect. Standard deviations are listed in parentheses.

Speed

Average WPMs over all 10 sessions were: 3-key 9.08 (1.31), 5-key 10.62 (2.61), 4-key 12.50 (3.91), 4-key-noviz 12.94 (3.99). By session 10 these improved to: 3-key 9.81 (1.33), 5-key 12.86 (2.26), 4-key 15.95 (3.22), 4-key-noviz 16.86 (3.06). Figure 5a depicts speeds over sessions (next page).

A main effect for WPM is significant over 10 sessions for *Session* ($F_{1,164}$=350.63, p<.01), *Method* ($F_{3,164}$=72.61, p<.01), and *Session* × *Method* ($F_{3,164}$=27.91, p<.01). That is, subjects sped up over time and did so at different rates with each method. Contrasts show that the speeds of the 4-key methods were not detectably different ($F_{1,164}$=1.45, ns), but were significantly faster than 3-key ($F_{1,164}$=202.78, p<.01) and 5-key ($F_{1,164}$=67.21, p<.01). Also, 5-key was faster than 3-key ($F_{1,164}$=26.88, p<.01).

Learning curves show high correlations (Figure 5a). Subjects learned the 4-key methods quickly, overtaking the selection methods by session 2. Subject 3 performed the fastest with all four methods (Figure 5b). Over 10 sessions, his WPM reached: 3-key 11.03 (1.35), 5-key 15.99 (1.65), 4-key 20.75 (1.47), 4-key-noviz 21.73 (1.50). Subject 3's single fastest phrase occurred in session 9 with the 4-key method at 24.06 WPM and 0% errors. This is remarkable for typing on four keys after ~90 minutes of practice.

Errors

Corrected errors are any letters backspaced during entry [10]. Thus, corrected errors reduce WPM. Corrected error rates (%) over 10 sessions were: 3-key 1.69 (1.61), 5-key 1.73 (1.49), 4-key 7.23 (4.26), 4-key-noviz 6.74 (4.30). A main effect for corrected errors is significant for *Session* ($F_{1,164}$=38.93, p<.01), *Method* ($F_{3,164}$=64.32, p<.01), and *Session* × *Method* ($F_{3,164}$=9.87, p<.01). Contrasts show the 4-key methods were not detectably different ($F_{1,164}$=0.75, ns), but were more error prone than 3-key ($F_{1,164}$=130.82, p<.01) and 5-key ($F_{1,164}$=125.69, p<.01). Also, 3-key and 5-key were not significantly different ($F_{1,164}$=0.02, ns). Although more error prone overall, 4-key errors dropped significantly over sessions ($F_{1,80}$=47.43, p<.01) and may have continued to do so over future sessions (Figure 5c).

Uncorrected errors are those that remain in the transcription [10]. Uncorrected error rates (%) over 10 sessions were: 3-key 0.60 (0.75), 5-key 0.96 (1.10), 4-key 1.69 (1.71), 4-key-noviz 2.00 (2.05). A main effect for uncorrected errors is significant for *Session* ($F_{1,164}$=4.00, p<.05) and *Method* ($F_{3,164}$=9.05, p<.01) but not for *Session* × *Method* ($F_{3,164}$=0.43, ns). Contrasts show that the 4-key methods were not detectably different ($F_{1,164}$=1.37, ns), but left more errors than 3-key ($F_{1,164}$=22.71, p<.01) and 5-key ($F_{1,164}$=10.42, p<.01). Also, the 3-key and 5-key methods were not significantly different ($F_{1,164}$=1.73, ns).

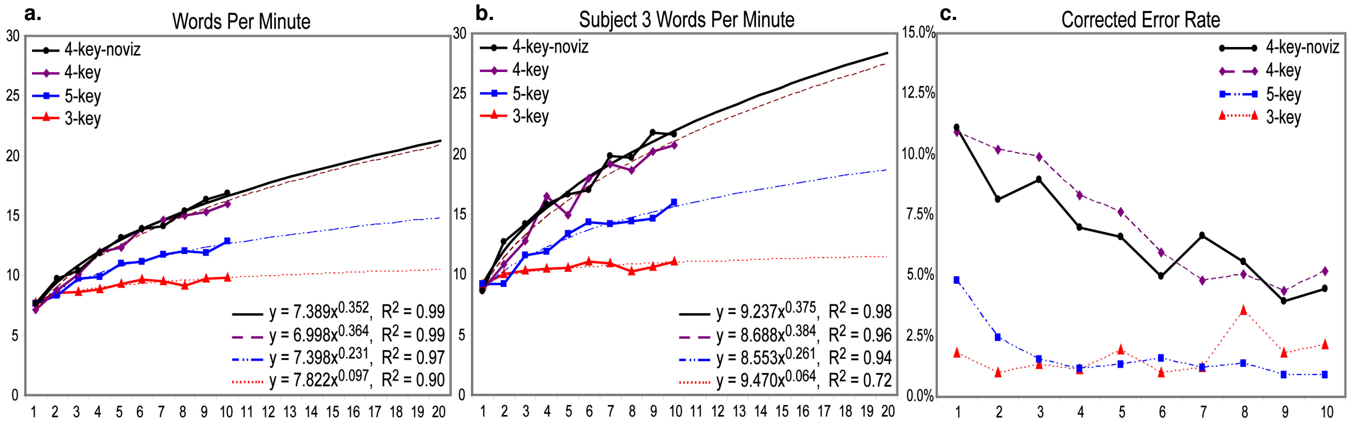

Figure 5. (a) Mean WPM and learning curves. **(b)** Fastest subject WPM and learning curves. **(c)** Corrected errors drop significantly for 4-key.

Clearly, speed and *uncorrected* errors are at odds—one can go faster if one leaves more errors. We can therefore use

$$AdjWPM = WPM \times (1 - Uncorrected\ Error\ Rate) \qquad (2)$$

as a unified measure. Means over 10 sessions were: 3-key 9.02 (1.29), 5-key 10.52 (2.58), 4-key 12.28 (3.79), 4-key-noviz 12.67 (3.87). By the 10th session these improved to: 3-key 9.74 (1.32), 5-key 12.73 (2.25), 4-key 15.70 (3.29), 4-key-noviz 16.42 (3.27). Since the uncorrected error rates of all four methods were low (≤2%), the 4-key methods were still significantly faster using AdjWPM than 3-key ($F_{1,164}$=195.42, p<.01) and 5-key ($F_{1,164}$=62.83, p<.01).

DISCUSSION

Despite being gestural, the 4-key methods were quickly learned, demonstrating the benefits of mnemonic gestures [12]. It is clear from the learning curves that more sessions, or sessions offering more practice, are needed to find the asymptotic speeds of the 4-key methods.

It is not surprising that the gestural methods were less accurate, although uncorrected error rates were fairly low for all four methods (≤2%). Other studies of selection vs. gestural methods [5] confirm the latter as often more error prone. For users to enter an error with a selection method, they must place the selector over the wrong letter and *still* choose to select it. In a gestural method, even a correctly-intended stroke may be poorly executed or misrecognized. Future work will improve accuracy by intelligently handling premature segmentations and by preventing backspaces for a brief period after non-recognitions.

CONCLUSION

We have presented a new 4-key text entry method based on mnemonic gestures. The method demonstrates how Roman letters can be mapped to four keys to create a quickly learnable gestural technique. Our new method is faster than previous selection-based few-key methods, although error rates were higher. Our results indicate that removing a visualization of the letter stroke being made is *not* detrimental to the gestural method, which may have positive implications for text entry in mobile, wearable, and assistive contexts.

ACKNOWLEDGEMENTS

This work was supported by Microsoft, General Motors, and the National Science Foundation under grant UA-0308065. Any opinions, findings, conclusions, or recommendations expressed in this material are those of the authors and do not necessarily reflect those of the National Science Foundation.

REFERENCES

1. Bellman, T. and MacKenzie, I.S. (1998) A probabilistic character layout strategy for mobile text entry. *Proc. Graphics Interface '98.* Canadian Info. Proc. Soc., 168-176.
2. Evreinova, T., Evreinov, G. and Raisamo, R. (2004) Four-key text entry for physically challenged people. *Adjunct Proc. UI4ALL '04.* http://ui4all.ics.forth.gr/workshop2004/
3. Isokoski, P. and Raisamo, R. (2000) Device independent text input: A rationale and an example. *Proc. ACM AVI '00*, 76-83.
4. Jannotti, J. (2002) Iconic text entry using a numeric keypad. Unpublished. http://www.jannotti.com/papers/iconic-uist02.pdf
5. Költringer, T. and Grechenig, T. (2004) Comparing the immediate usability of Graffiti 2 and virtual keyboard. *Extended Abstracts ACM CHI '04*, 1175-1178.
6. MacKenzie, I.S. (2002) KSPC (keystrokes per character) as a characteristic of text entry techniques. *Proc. Mobile HCI '02.* Springer-Verlag, 195-210.
7. MacKenzie, I.S. (2002) Mobile text entry using three keys. *Proc. ACM NordiCHI '02*, 27-34.
8. MacKenzie, I.S. and Soukoreff, R.W. (2003) Phrase sets for evaluating text entry techniques. *Extended Abstracts ACM CHI '03*, 754-755.
9. Orth, M., Post, R. and Cooper, E. (1998) Fabric computing interfaces. *Conference Summary ACM CHI '98*, 331-332.
10. Soukoreff, R.W. and MacKenzie, I.S. (2003) Metrics for text entry research: An evaluation of MSD and KSPC, and a new unified error metric. *Proc. ACM CHI '03*, 113-120.
11. Soukoreff, R.W. and MacKenzie, I.S. (1995) Theoretical upper and lower bounds on typing speed using a stylus and soft keyboard. *Behavior and Information Technology 14* (6), 370-379.
12. Wobbrock, J. O., Aung, H. H., Rothrock, B. and Myers, B. A. (2005) Maximizing the guessability of symbolic input. *Extended Abstracts ACM CHI '05*, 1869-1872.
13. Wobbrock, J.O., Myers, B.A. and Kembel, J.A. (2003) EdgeWrite: A stylus-based text entry method designed for high accuracy and stability of motion. *Proc. ACM UIST '03*, 61-70.