

Comparing performance with two methods of text input with a handheld keypad: Four-way vs. Multi-press

Leonid Papkov and Prof. Joel Norman

Ergonomics and Human Factors Unit

Institute of Information Processing and Decision Making

The University of Haifa, Israel

July 2003



אוניברסיטת חיפה, טל. 8249430-04, פקס. 8249431-04, דואר אלקטרוני etti.l@psy.haifa.ac.il
UNIVERSITY OF HAIFA, HAIFA 31905, ISRAEL, TEL. 972-4-8249430, FAX. 972-4-8249431, e-mail: etti.l@psy.haifa.ac.il

EXECUTIVE SUMMARY

The present study compared two methods of text entry speed using handheld keypads; the conventional Multi-press (SMS) method and a novel Four-way method developed by JetWay Technologies. This new method utilizes specialized keys that can be pressed in one of four directions, allowing a one-press character entry.

Ten students participated in the experiment for pay, including bonuses for fast and accurate performance. After an initial test and training session each participant took part in nine experimental sessions; five participants used the Four-way method and five used the Multi-press method. Overall, each participant entered a total of 450 4-7-word sentences (mean of 30.5 characters per sentence).

Both speed and accuracy of performance were analyzed, but as accuracy was very high throughout, speed served as the primary measure of performance. The four-way method yielded higher speeds throughout, even at the beginning of the experiment. This superiority increased with practice, indicating that for expert users the advantage of the Four-way method would be even greater than that found in the last sessions. The results showed linear improvement with practice for both methods indicating that the participants were far from reaching asymptotic performance.

It was suggested that the Four-way method yields superior results because it does not require multiple presses for a single character, and no time-out pauses between entering successive characters using the same key. The speeds attained in the present study using the Four-way method were also faster than those reported in a study of another single-press method, Fastap, but such comparisons are only suggestive.

All in all, the results of the present study clearly point to the superiority of the Four-way method over the Multi-press method. What is more, they indicate that this difference is likely to become even larger with further practice. Comparing the results of this study with other similar studies also hints at the possible superiority of the Four-way method over other newly developed devices for handheld text entry.

Text entry using handheld keypads has become extremely popular of late. This is mainly due to the use of “short message service” (SMS) in cell-phone communication. It has been estimated that currently 24 billion SMS messages are sent every month (Cockburn & Siresena, 2003). However, there are also other uses for handheld keypads, such as entering messages in interactive television services.

Most commonly SMS utilizes the Multi-press (also called Multi-tap) input technique. This consists of multiple presses of one of 12 keys to signify different letters or numerals. This is a cumbersome method where some letters require three key-presses, and the high-frequency letter “s” requires four presses, while numerals require four to five presses. Another source of this method’s inefficiency is the fact that a pause (“time-out”) is required before the same key can be used to enter another character.

To overcome these obvious deficiencies of the Multi-press method alternative methods have been suggested. Two of the more prominent are the T9 (Silfverberg, MacKenzie, and Korhonen, 2000) and Fastap (Sirisena, 2002) methods. Briefly, T9 is a linguistically based approach, where the more frequent words in a language are automatically presented on the basis of the key-sequence pressed. Words not in the dictionary have to be tapped in using the Multi-press method, and when the device’s first “guess” is incorrect (about 7%) the user has to select the word from a list that is provided. The Fastap method is based on a different kind of keypad consisting of 27 keys for the 26 letters and a space bar. Other characters such as numerals can be entered by pressing the four keys surrounding the numeral (“chording”). Studies of text entry with these two new methods have indicated that they are superior to Multi-press (see Discussion).



Figure 1. The JetWand keypad. (Sketch of layout on the left and photograph of the keypad on its cradle on the right. Note the infrared communication system on the bottom of the cradle and on the top of the keypad).

Another alternative approach to handheld keypad text entry has recently been introduced by JetWay Technologies, utilizing a Four-way key press technique. Specially designed four-way keys can be pressed in one of four directions, with each direction yielding a different character. The current version, JetWand (see Figure 1) is larger in size than mobile phones (~20 cm long) and is similar in size to a remote TV control. In the present study we only utilized the 12 central keys and the space bar. The front panel includes many other function keys that were not used in the present study.

In this report we detail a study aimed at comparing this new Four-way method with the older Multi-press method in a text-entry task. Ten paid participants, five with each method, underwent nine experimental sessions during which they entered a total of 450 sentences. The speed and accuracy of their performance were analyzed and are presented below.

Method

Participants. Ten University of Haifa students (3 men and 7 women) with good knowledge of English were selected from among 15 applicants. The other five were rejected after performing poorly on a test of English spelling. The participants were paid a base fee for participating, and a bonus based on how well they performed.

Apparatus. The input device was a handheld keypad supplied by Jetway Technologies. This keypad allowed pressing each key in one of four directions (Four-way), each direction for a different character. The computer interface software also allowed using the same keypad as a Multi-press keypad, where the number of taps indicated the required character (as in the SMS cellphone configuration).

The experiment was carried out on a Compaq EY Pentium III 733 MHz PC with a 17" MAG screen. The participants wore HD 1010 stereo headphones throughout the experiment.

Materials. A bank of 150 English sentences was constructed. These contained between 4 and 7 words, with the number of characters (including spaces and full stop at end) ranging between 25 and 39. The sentences were selected from English conversation primers (Brinton et al., 1972; Sutherland, 1980; White, 1979), and included few difficult words. These sentences were recorded and presented both

in a visual and auditory mode in the main part of the experiment. Only lower-case letters were used.

Procedure. The experiment consisted of 10 sessions. In the first session the participant's English spelling was tested, as well as their skill in Multi-press keying. The test of English spelling consisted of the typing of twenty of the apparently more difficult sentences in the bank, using the regular QWERTY keyboard. The sentences were presented only through the headphones. The criterion for continuing in the rest of the experiment was fewer than 40 spelling errors on the 20 sentences. This was followed by the participants' first encounter with the handheld keypad, where they had to type 30 random four-digit numbers that appeared on the screen. This served as preliminary training with the new device and a test of their experience with similar keypads. This was followed by the typing of 20 easy sentences using the Multi-press method. The results of the latter were used to assign the ten participants to one of the two keying methods, Multi-press or Four-way (see Table 1).

Participant No.	Gender	CPS	EPS	Overall ranking	Method assigned
3	F	0.92	1.00	3	Multi-press
5	F	0.45	1.00	7	Multi-press
6	F	0.67	0.75	4	Four-way
8	F	0.87	1.60	5	Multi-press
9	M	0.78	0.80	2	Four-way
10	M	0.47	1.35	9	Multi-press
12	M	0.54	1.00	6	Four-way
13	F	0.95	0.65	1	Multi-press
14	F	0.53	1.55	8	Four-way
15	F	0.62	2.30	10	Four-way

Table 1. Results of the participants in the 1st session (20 sentences typed using the Multi-press method), including the overall ranking based on the speed and accuracy of the performance and used for assignment to experimental groups.

The following nine experimental sessions were identical for each participant, five of whom used Multi-press throughout, and the other five Four-way. In each session the participants were presented with five blocks of 10 sentences. These were presented both auditorily and visually and disappeared before the participant could start typing. Error corrections were not allowed during the experiment. The participants were instructed to work as quickly and as accurately as they could, and were informed that the monetary bonus would depend on both factors. At the end of each block the participants received feedback about the average speed of typing, the number of errors, and the bonus accrued. The entire bank of 150 sentences (“set” in the results section below) was presented in random order in each three consecutive sessions. Thus, the participants encountered each sentence three times during the entire experiment.

Results

Two measures of performance served in the analysis of the results. These were Characters per Second (**CPS**)¹ and Errors per Sentence (**EPS**). The raw overall results per session of CPS are presented in Figure 2, and those for EPS in Figure 3. As can be seen in both figures the results are not entirely monotonic. This is most probably due to the fact that single sessions were not equal in that they contained different sentences. Therefore, the following analyses will all be on the three sets of data. These are three sets of three consecutive sessions that consist of the same 150 sentences.

¹ We have chosen to report the entry speed in CPS rather than the more common Word per Minute (WPM) as we discovered that different research groups define WPM differently. Some define a word as 5 characters including a space while others use 5.98. To transform CPS to the former simply multiply by 12, and to the latter by approximately 10.

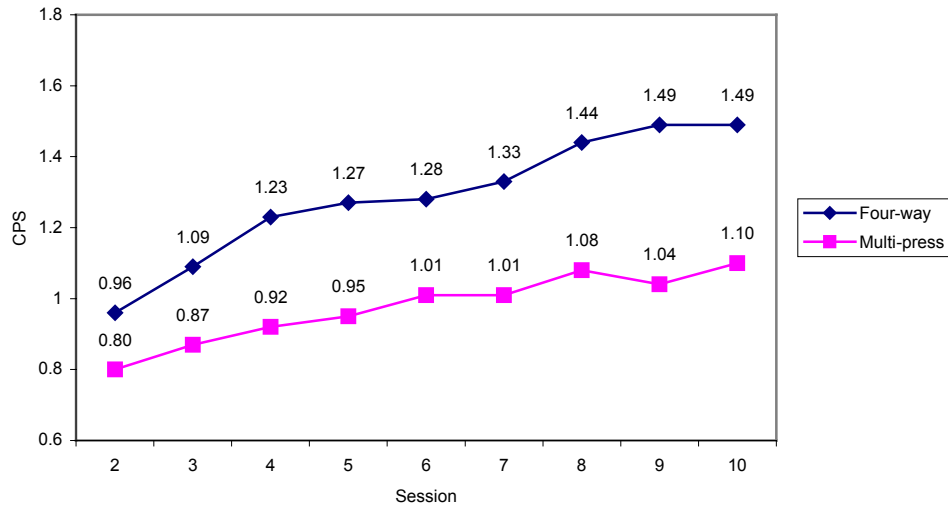


Figure 2. Mean text entry speed (CPS) over the nine experimental sessions.

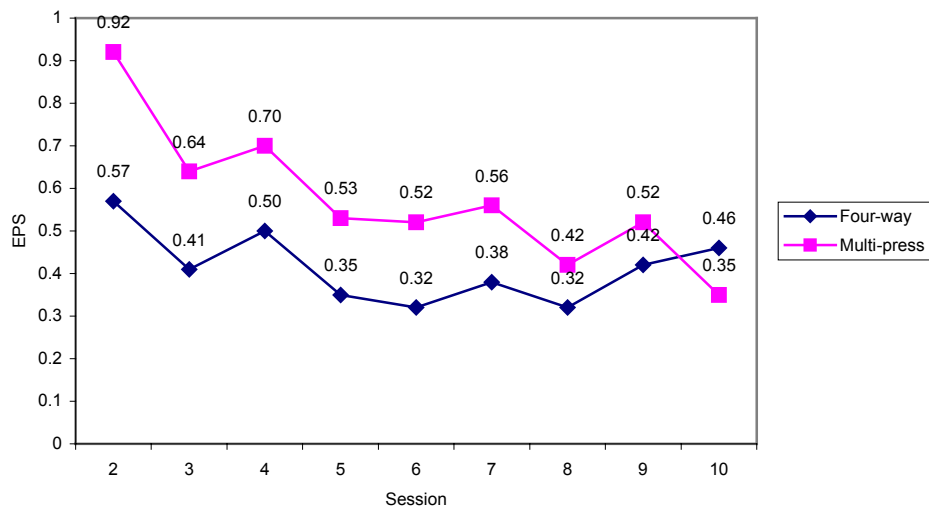


Figure 3. Mean errors per sentence (EPS) over the nine experimental sessions.

From the raw data presented in Figures 2 and 3 one can see that overall performance with the Four-way method is superior to that of the Multi-press method. This is also very clear in the analyses over complete data sets. In Figure 4 we see the overall superiority of the Four-way method on text entry speed. These data were analyzed using a mixed design ANOVA. There was an overall main effect of method,

$F(1,8)=6.36$, $p<0.036$, where Four-way was superior throughout (overall means: Four-way 1.29 CPS; Multi-press 0.92 CPS). As expected, there was a significant improvement over the three sets, $F(2,16)=49.66$, $p<0.0001$. In Figure 4 one can see an interaction, where there is greater improvement over sets for the Four-way than the Multi-press method. This interaction was found to be statistically significant, $F(2,16)=7.16$, $p<0.006$.

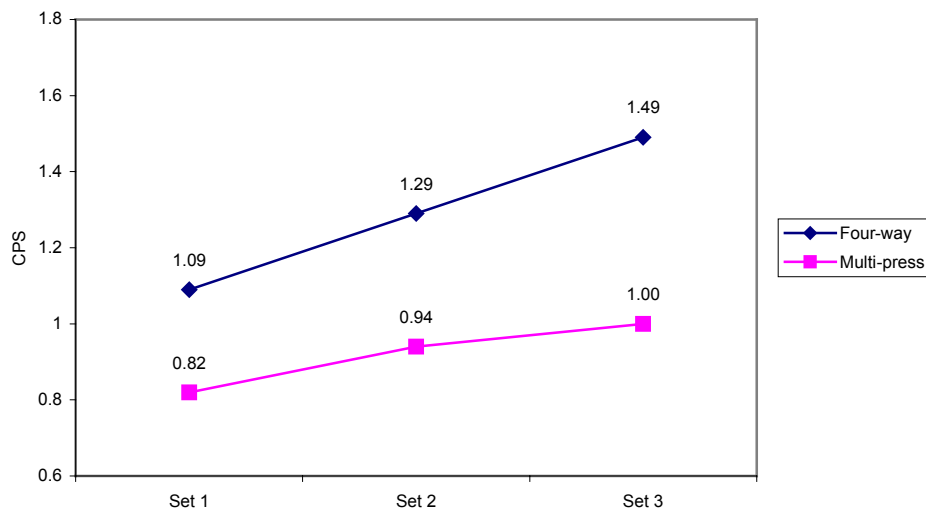


Figure 4. Mean text entry speed (CPS) over the three sentence sets.

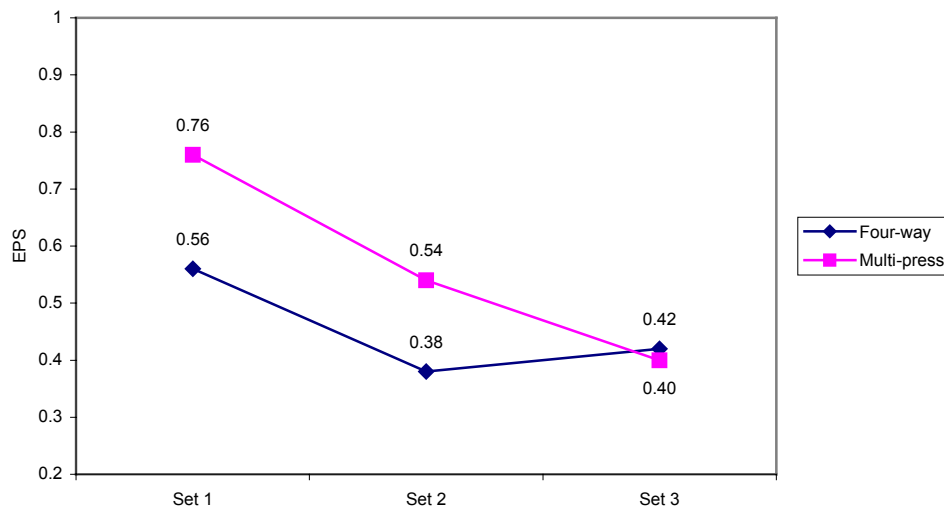


Figure 5. Mean errors per sentence (EPS) over the three sentence sets.

A similar analysis of EPS (see Figure 5) yielded a non-significant main effect of method, $F(1, 8)=0.99$, $p<0.350$, in spite of the apparent superiority of the Four-way method (overall means: Four-way 0.45 EPS; Multi-press 0.56 EPS). On the other hand, there was a near significant decline in errors with practice, $F(2,16)=3.58$, $p<0.052$. While there is an apparent interaction, it does not attain statistical significance, $F(2,16)=0.60$, $p<0.562$.

Some further analyses were carried out. The Multi-press method requires a “time-out” before the user can type another letter using the same key. We divided the 150-sentence bank into 6 groups according to the number of time-outs in a given sentence. Figure 6 presents typing speed as a function of the number of time-outs in the sentences. As expected, the Four-way method is not affected by this variable, but as can be seen there is a steady decline in CPS with increasing number of time-outs for the Multi-press method. This interaction was found to be statistically significant, $F(5,40)=12.72$, $p<0.0001$.

An additional analysis examined improvement within session over the subsequent five blocks. This analysis indicated no such improvement. Improvement only occurred over sessions.

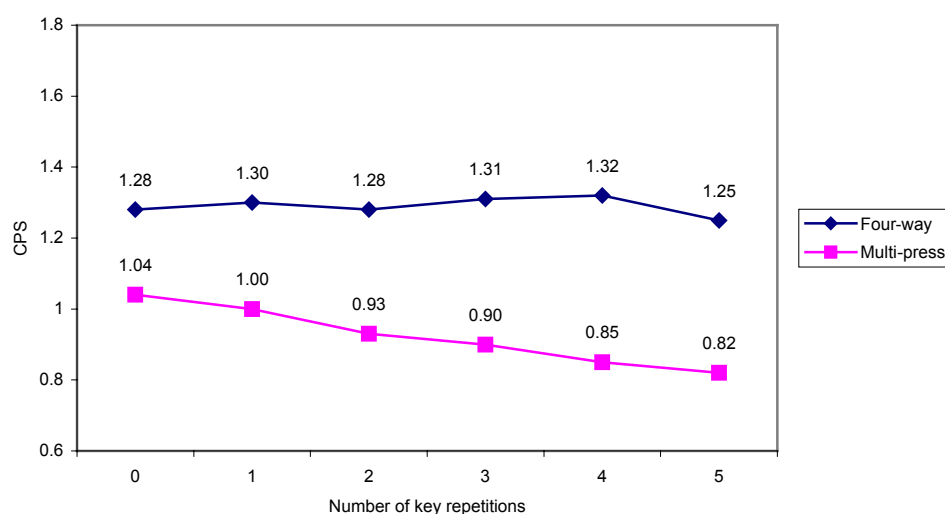


Figure 6. Mean text entry speed (CPS) as a function of key repetitions (see text).

An informal look at the ten participants' individual results was carried out. Figure 7 presents the results for all participants over the three sets. The overall superiority of the Four-way method can be seen in the data of three of the five participants who used this method, while a fourth participant was not better than the best users of the Multi-press method, and a fifth participant performed more poorly than three of the Multi-press group.

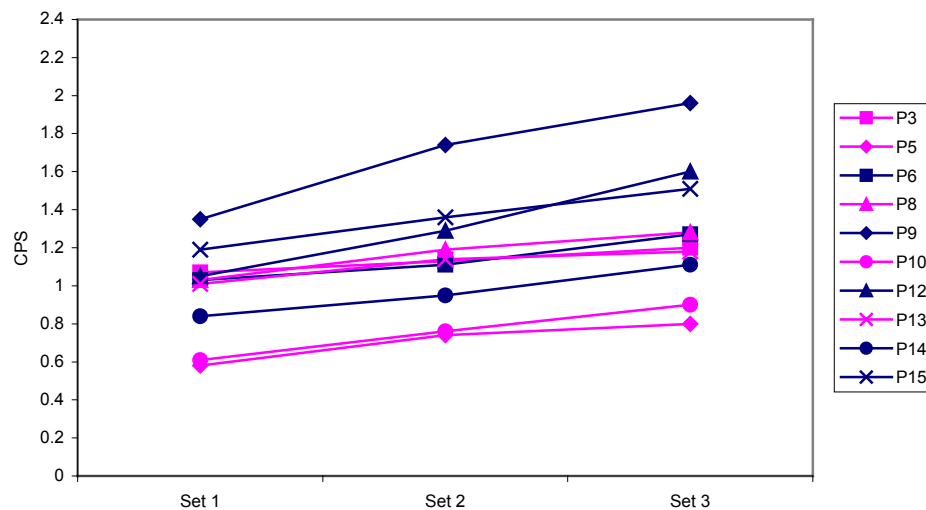


Figure 7. Mean text entry speed (CPS) for individual participants over the three sentence sets. Participants using Multi-press appear in pink; those using Four-way in blue.

Discussion

The results of the present study point to the superiority of the Four-way method of text entry over that of the Multi-press method. This is true for both the speed of text entry (CPS) and the accuracy measure (EPS). However, it should be noted that the participants knew that the bonus would be reduced for errors in typing, and so they worked very accurately, with the initial error rates not exceeding 1 error per sentence and dropping to less than 0.5 errors per sentence towards the end of the experiment. Therefore, the speed measure, CPS, is the primary measure of interest in comparing the two methods.

The main results are depicted in Figure 4, where it can be seen that not only is the Four-way method the superior of the two, but it also shows greater improvement with practice than the Multi-press method. The graphs for both methods are quite linear over the three sets, apparently indicating that with further practice the user will become even faster. Looking at the results for individual participants, it can be seen that fastest (by far) participant (No. 9 using Four-way) also yields a linear increase with practice, indicating that the results of the present study are still far from a performance ceiling. Other studies indicate that much more practice is needed to achieve asymptotic performance. For example, MacKenzie and Soukoreff (2002) report a study where participants took part in twenty 20-minute sessions, and the results did not reach peak performance. They further extrapolate the practice effects and their graph suggests that for a newly learned keyboard asymptotic performance is not attained even after 50 sessions!

It should be further noted that several of the participants reported that they had quite a lot of practice using SMS cell-phone communication, so that some of those tested with the Multi-press method had a “head start” in comparison to the Four-way method, where none had any previous experience.

One clear inherent advantage of the Four-way method is its not requiring a time-out between successive letters (key repetitions) entered on the same key (e.g. “p” before “r”) as in the Multi-press method. Indeed, in Figure 6 it was seen that there was no effect of the number of key repetitions on Four-way performance, while there was a clear decline in entry speed in the Multi-press method. It should be noted, however, that even when there were no letter repetitions in a sentence the Four-way method was still the faster of the two. This is most probably due to the fact that unlike the Multi-press method, it does not require more than one press per character.

It is difficult to compare our results with those of related studies for several reasons: The exact amount of practice differs between studies; the text materials are different; and the keypads used are different. In general it would appear that the participants in our studies achieved greater speeds than in most of the previous studies. Even comparing the results in our study on Multi-press yields faster entry speed than other studies. For example, Butts and Cockburn (2001) report a mean speed of 0.64 CPS, while we find a mean of 0.92 CPS. The reasons for this difference are not clear, but it is probably due to the larger keypad used in the present study and to the fact that the participants were motivated (bonus!) to type quickly and had more practice. How does the faster Four-way method compare with reported results in the other system using single key-presses, Fastap? In a study of entry speed using several methods including Fastap, Cockburn and Siresena (2003) report a mean input speed of 8.5 WPM, equivalent to 0.85 CPS. This is quite a bit slower than the mean attained by the participants using the Four-way method in the present study, 1.29 CPS. Speculating about the reason for the superiority of the Four-way method, one might attribute it to the fact that it only requires 13 keys to be pressed, while Fastap requires 27 keys.

In sum, the present study clearly pointed to the superiority of the Four-way method over the Multi-press method, also indicating that this difference is likely to become even larger with extended practice. What is more, while comparisons between studies using different materials and devices are not well founded, it would appear that the Four-way method is also superior to the other newly developed handheld text entry methods.

References

- Brinton, E., Davies, P., Plumb, W., White, C., & Shepherd, J. (1972). Active Context English: Book Three. London: Macmillan Press.
- Butts, L., & Cockburn, A. (2001). An evaluation of mobile phone text input methods. Proceedings of Australian User Interface Conference, 2002, pp. 55-59.
- Cockburn, A., & Siresena, A. (2003). Evaluating mobile text entry with the Fastap keypad. People and Computers XVII (Volume 2): British Computer Society Conference on Human Computer Interaction. In Press.
- MacKenzie, I. S., & Soukoreff, R. W. (2002). Text entry for mobile computing: Models and methods, theory and practice. Human-Computer Interaction, 17, 147-198.
- Silfverberg, M., MacKenzie, I. S., & Korhonen, P. (2000). Predicting text entry speed on mobile phones. Proceedings of the ACM Conference on Human Factors in Computing Systems - CHI 2000. New York: ACM, pp. 9-16.
- Sirisena, A. (2002). Mobile Text Entry. Retrieved July 27, 2003, from http://www.cosc.canterbury.ac.nz/research/reports/HonsReps/2002/hons_0201.pdf.
- Sutherland, K. (1980). English Alpha. Boston: Houghton Mifflin.
- White, R. V. (1979). Functional English: Consolidation. Sundbury-on-Thames: Nelson.