An Alternative to Push, Press, and Tap-tap-tap: Gesturing on an Isometric Joystick for Mobile Phone Text Entry

Jacob O. Wobbrock,[1,2] Duen Horng Chau[2] and Brad A. Myers[2]

[1]The Information School
University of Washington
Mary Gates Hall, Box 352840
Seattle, WA 98195-2840
wobbrock@u.washington.edu

[2]Human-Computer Interaction Institute
School of Computer Science
Carnegie Mellon University
Pittsburgh, PA 15213 USA
⟨dchau, bam⟩@cs.cmu.edu

ABSTRACT
A gestural text entry method for mobile is presented. Unlike most mobile phone text entry methods, which rely on repeatedly pressing buttons, our gestural method uses an isometric joystick and the EdgeWrite alphabet to allow users to write by making letter-like "pressure strokes." In a 15-session study comparing character-level EdgeWrite to Multitap, subjects' speeds were statistically indistinguishable, reaching about 10 WPM. In a second 15-session study comparing word-level EdgeWrite to T9, the same subjects were again statistically indistinguishable, reaching about 16 WPM. Uncorrected errors were low, around 1% or less for each method. In addition, subjective results favored EdgeWrite. Overall, results indicate that our isometric joystick-based method is highly competitive with two commercial keypad-based methods, opening the way for keypad-less designs and text entry on tiny devices. Additional results showed that a joystick on the back could be used at about 70% of the speed of the front, and the front joystick could be used eyes-free at about 80% of the speed of normal use.

Author Keywords
Text input, gestures, unistrokes, isometric joysticks, mobile phones, pointing, Multitap, T9, EdgeWrite, Smartphones.

ACM Classification Keywords
H5.2. Information interfaces and presentation: User interfaces — *Input devices and strategies*.

INTRODUCTION
Among computing platforms, mobile phones are easily the most widespread. At the close of 2003, there were 400 million mobile phone users in Europe, 150 million in the United States, and 270 million in China with an additional 5 million new users each month. Phones are becoming increasingly powerful, capable of displaying full-color graphics, animations, and quality sound. We are rapidly approaching an era in which mobile phones are serving as streaming media players, pocket jukeboxes, digital cameras, portable internet terminals, email clients, handheld televisions, pocket maps, and video game platforms [24]. In addition, phones may become people's personal remote controls for accessing computers in the environment [15].

Figure 1. (a) Our Red•E SC1100 Smartphone with an IBM Track-Point isometric joystick embedded in the front. The phone's screen shows our text entry software. **(b)** The same phone showing the other joystick embedded in the back.

However, this vision is compromised by the poor input methods available on today's phones [6]. Phone keypads support only discrete keypresses and are incapable of expressing the kind of fluid control essential to many new phone-based applications. For example, video games, music or video playback, scrolling long song lists, and panning maps or photos will all require fluid control. Although joysticks traditionally can provide such control, the joysticks found on today's phones are really just four-way switches that cannot provide a range of inputs. We therefore envision a future in which phones may be equipped with more expressive joysticks, such as isometric joysticks.

This research anticipates this future by studying how isometric joysticks can be used to address a particularly lacking aspect of mobile phone use—that of text entry [1]. Given that future phones may have isometric joysticks (or similar), it is worth investigating how these devices may be used for more natural or expressive input. Specifically, this research investigates whether an isometric joystick can be competitive with predominant keypad-based methods for mobile phone text entry. Using a physical phone prototype we created with an embedded isometric joystick (Figure 1a), we compared a gestural text entry method that uses the *EdgeWrite* alphabet [26] to *Multitap* and *T9* (www.tegic.com) in a longitudinal study over 30 short sessions. Both character-level and word-level [29] versions of EdgeWrite are involved in the study.

We also compare a back-mounted isometric joystick for use with the index finger (Figure 1b) to the traditional front-mounted joystick for use with the thumb, finding the former to be about 70% as fast as the latter for text entry. We also include results indicating that EdgeWrite is about 2.5 times faster than Multitap when the phone cannot be seen, which has positive implications for eyes-free mobile use.

Overall, our study results show that EdgeWrite text entry on an isometric joystick is highly competitive with both Multitap and T9, and that subjects generally preferred EdgeWrite. Advantages of our design include the ability to write more "by feel" rather than by sight, and the ability to enter text in a very small space. This work thus paves the way for keypadless designs and text entry on tiny devices.

RELATED WORK

An *isometric joystick* is a joystick that senses force but does not move. Isometric joysticks were part of the early formal study of input devices by Card *et al.* [2] that helped extend Fitts' law to pointing devices. Thereafter, isometric joysticks became popular with the advent of the *IBM TrackPoint*, the small "eraser head" found on ThinkPad laptops that replaces the mouse [18]. However, Mithal and Douglas [13] found that cursor control with an isometric joystick involves many more fine sub-movement corrections than with a mouse, which is partly why isometric joysticks sometimes feel less accurate.

TrackPoint isometric joysticks have been embedded in research prototypes before, although none were mobile phones. Zhai *et al.* [33] combined a joystick with a desktop mouse (Figure 2a) to facilitate pointing with the mouse and scrolling with the finger-controlled joystick. They found that this design was faster and better liked than the IntelliMouse scroll wheel. Silfverberg *et al.* [20] embedded an isometric joystick in two handheld prototypes for pointing. They found that their two-handed prototype could achieve about 78% of a ThinkPad joystick's throughput—and their one-handed prototype (Figure 2b) about 68%—when using the thumb for control and a separate button for selection. Zaborowski [32] invented *ThumbTec* (Figure 2c), a thumb-controlled isometric joystick used in combination with three switches for playing music. Salem and Zhai [19] put a TrackPoint in a mouthpiece for tongue-controlled pointing, finding this to be feasible as an initial prototype.

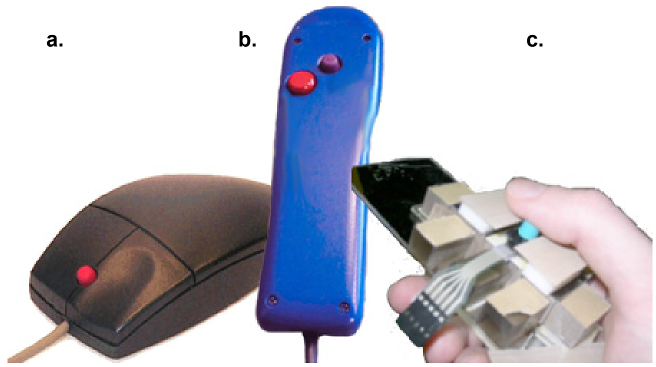

Figure 2. (a) The *Joystick Mouse* by Zhai *et al.* [33]. **(b)** The *one-handed pointer* by Silfverberg et al. [20]. **(c)** The *ThumbTec* musical input device by Zaborowski [32]. Images used with permission.

Although isometric joysticks have not been previously used for text entry, other types of joysticks have. For example, Wilson and Agrawala [25] used the two joysticks on a Microsoft Xbox game controller to select letters from two halves of a split on-screen keyboard. We have also used a game controller joystick to perform EdgeWrite gestures within a plastic square [27]. Importantly, this game controller design differed substantially from the current isometric joystick method in that it relied on the *absolute* position of the joystick within a physical square. The current method, in contrast, has no notion of absolute position and no physical square around its stick. Instead, it uses force vectors to construct EdgeWrite gestures. Other displacement joystick-based text entry methods include *weegie* [4], *Quikwriting* [7], and *Xnav* [9], all of which use some form of menus or regions from which letters are selected.

PROTOTYPE DESIGN AND IMPLEMENTATION

Our mobile phone hardware prototype [3] was built using a Red•E SC1100 Smartphone. In augmenting this device with an IBM TrackPoint isometric joystick, we made it a priority to retain as much of the built-in functionality of the Smartphone as possible. In particular, we made sure that the Smartphone's screen would still work. Unfortunately, the circuit for the phone's keypad had to be removed to make room for the isometric joystick circuit, so the keys themselves were not functional. Of course, the built-in four-way switch joystick that came with the Smartphone was also removed. The back-mounted joystick required the removal of a speaker from the back of the device. It also required that we drilled a hole in the device chassis. Hot glue was used to mount both the front and back joysticks securely.

The outputs from the two joysticks were attached to a long PS/2 cable that emerged from the bottom of the phone and connected to a desktop PC. The PC ran software that interpreted the user's actions with the joysticks and sent the results back to the phone. Software running on the phone displayed the stroke traces made by the user and any

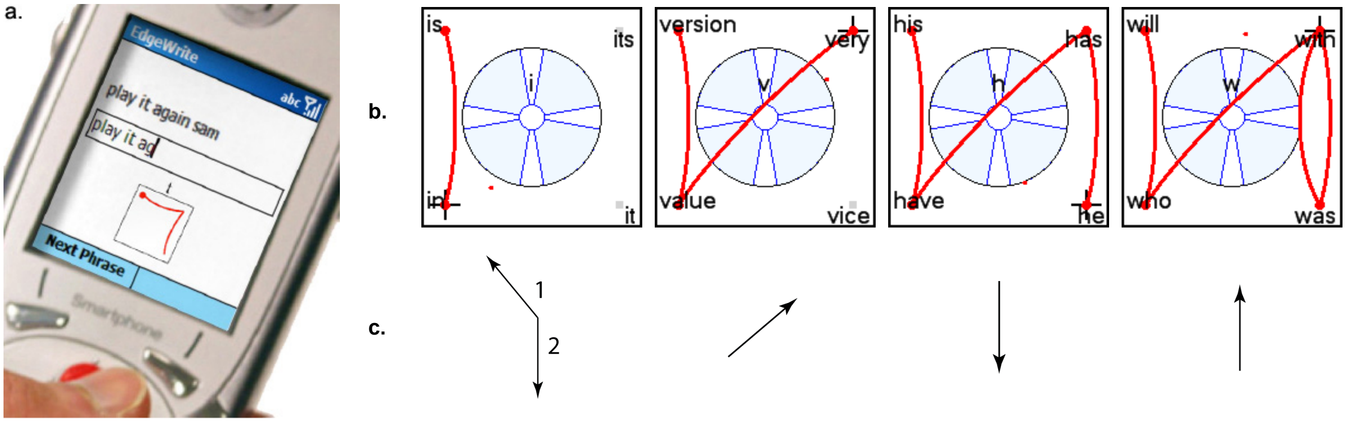

Figure 3. The software involved in the isometric joystick text entry system. **(a)** Smartphone software displaying EdgeWrite stroke traces and recognition results. In this image, the letter "t" is being made. **(b)** The desktop software showing a series of strokes that make the letter "w". Along the way, an "i", "v", and "h" are made, and English frequency-based word completions are presented. **(c)** The force vectors involved in making each stroke. The first stroke for "i" requires two force vectors: one to indicate the initial corner, and then one to move downward.

character-level or word-level stroke recognition results. Thus, although a PC was used in the loop for processing, subjects could attend to the phone screen as if no PC were used. Our prototype's software had two main components. One component ran on the Smartphone and was in charge of receiving data from the PC via TCP/IP over USB (Figure 3a). This data consisted of (1) EdgeWrite corner values, which fully define EdgeWrite strokes and were used in drawing traces; (2) any recognized characters or words, which were displayed in a text box; and (3) strings presented during user studies for subjects to transcribe.

The other main component was the software running on the PC. This software consisted of (1) the EdgeWrite stroke recognizer that interpreted isometric joystick movements and turned them into characters or words (Figure 3b); and (2) the user test software *TextTest* [30] that generated the presented strings [12] to be transcribed by subjects and logged their subsequent text entry activity.

The design of the EdgeWrite stroke recognizer for isometric joysticks was based on the one used for Trackball EdgeWrite [31]. This design could easily be adapted to isometric joysticks because trackballs and isometric joysticks have similar properties—e.g., neither device has a notion of position, but only a change in either rotation or force, and neither device actually translates when used.

In effect, the underlying writing design interprets users' *force vectors* as indicating one of four corners of an imaginary EdgeWrite square in which letters are made (Figure 3c). EdgeWrite's straight line-based alphabet enables users to avoid having to make the kinds of smooth curvy gestures that would be required for Graffiti 2 or natural handwriting. Importantly, our subjects indicated that using EdgeWrite still felt like writing hand-printed letters because of the Roman-like designs of the EdgeWrite characters [26]. Segmentation between letters is achieved when users briefly cease their force on the joystick. Readers interested in further details of the writing mechanics are directed to the prior work on Trackball EdgeWrite [31].

Character-level and word-level entry coexist in EdgeWrite without interference. For word-level stroking [29], word-completions are presented based on the current prefix and stroke underway (Figure 3b). The presence of words at the four corners of the EdgeWrite square does not mean that users must utilize them. Instead, users are free to stroke character-by-character without ever selecting a word. To select a word, users simply make a stroke into the corner of the desired word and then segment that stroke as usual by briefly ceasing their force on the isometric joystick. All EdgeWrite letters require at least two corners to make, so there is no conflict between selecting a word and stroking a character. Erroneous completions can be corrected with a single backspace from right-to-left along the *bottom* of the square. This removes the completed suffix and restores the prior completions. In contrast, a normal single-character backspace is made across the *top* of the square. Thus, the "w" stroke in Figure 3b would produce only the letter "w" despite the presence of "w"-words. However, after that "w" is entered, a stroke to the upper-right corner would enter "ith" to complete the word "with". Importantly, words are always shown in the same corners, allowing for improvement through memorization and motor repetition.

Although our software for the isometric joystick was based on Trackball EdgeWrite, modifications were crucial in getting the writing to "feel right" on the phone. For example, the sensitivity of the joystick, the size of the virtual EdgeWrite square, the segmentation timeout, and the angular regions in which users' force vectors are interpreted were all adjusted differently from Trackball EdgeWrite. Most of these adjustments were made in ways that lessened the force required to write comfortably with the joystick.

Concerning the back-mounted joystick and its use with the index finger (Figure 4a), we were at first unsure whether users would prefer to write in a visually-oriented manner, as if looking *through* the device, or in a motor-oriented manner, maintaining the same finger path they would use on the front of the device (and hence, visually mirrored when looking through the device). Figure 4b illustrates this issue

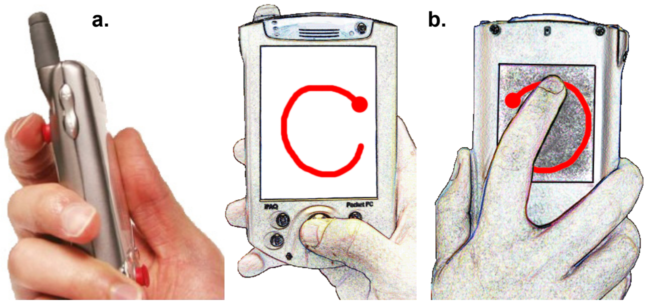

Figure 4. (a) The back joystick being used with the index finger. **(b)** This trace of "C" is visually correct from the front but motor-reversed from the perspective of the index finger on the back.

for a hypothetical PDA with a touchpad on its back.[1]

In a pilot study, we found that subjects preferred to write in a visually-oriented manner as shown in Figure 4b. This finding is consistent with prior work for letters made in different orientations on the human body [16]. Because of this, we had to add software to *reflect* the horizontal movement of the joystick. We also found that users' index fingers operated the isometric joystick at an angle. This meant that what felt like pushing "straight up" (90°) was actually at an angle (110°-130°). Thus, we implemented the ability to arbitrarily reflect and rotate the input space (Figure 5). Through a brief period of experimentation, users could discover the proper settings that resulted in accurate and comfortable movements.

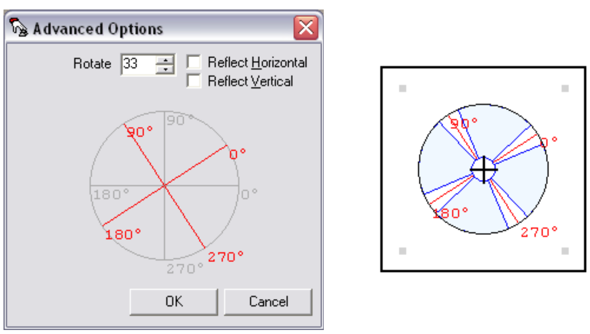

Figure 5. The configuration dialog and illustrated result for arbitrary rotations and reflections of the isometric joystick input space.

EVALUATION

To assess using isometric joysticks with gestures for mobile phone text entry, we conducted a study comparing our prototype to Multitap and T9, which respectively account for 32.3% and 10.5% of phone text entry use among a surveyed population mostly in the U.S. in August 2005 [10].

Method

We conducted our study over multiple short sessions designed to simulate "daily intermittent use," which is typical with mobile devices.

[1] We illustrate this concept with a back-mounted touchpad so that a *finger path* can be shown, which is more difficult to illustrate with a finger on a back-mounted isometric joystick.

Subjects

Six paid subjects ages 21-28 were recruited from the local university community. Not all were students, and half were *not* technology majors. Four were male. Although all of the subjects owned mobile phones, their previous use of Multitap was very limited, involving only the entering of contact information. None were active users of text messaging, none had used T9, and none had ever used EdgeWrite. All six subjects participated in each part of the study.

Apparatus

For EdgeWrite, subjects used our isometric joystick phone prototype as described in the previous section. However, in pilot testing we found that latencies between the PC and phone were undesirable, making stroke feedback a little bit too slow. Therefore, subjects held the phone in a natural position adjacent to the PC monitor and observed their strokes and text input there. They were able to see their phone, hand, joystick, input, and output all at once, as the physical distance between these items was ≤3 inches.

As stated above, the joystick phone's keys were disabled since its keypad circuit had to be removed. Therefore, a second Smartphone, an i-Mate Smartphone 2 that was similar in size and shape to the joystick phone, was used for Multitap and T9. This phone ran the user test software and was connected to the desktop PC via TCP/IP over Bluetooth. Latency was not an issue since no stroke drawing was necessary. For both phones, the desktop PC logged text entry activity for later analysis.

Procedure

The evaluation was within-subjects and consisted of five parts in order to assess different aspects of our isometric joystick phone prototype. Part 1 consisted of 15 short sessions in which subjects entered phrases [12] of about 30 characters each using character-level EdgeWrite and Multitap. In each session, subjects warmed up with 2 phrases before testing on a set of 8 to 12 phrases, depending on the session number. In later sessions, subjects were faster so 12 phrases could be entered. By the end of the study, sessions only took about 20 minutes. Importantly, the number of phrases entered in a given session was *always* identical for EdgeWrite and Multitap, and measures were the average speeds and uncorrected error rates for the test phrases entered by subjects in that session.

Part 2 of the evaluation consisted of 15 additional sessions, this time comparing *word-level* EdgeWrite and T9. The T9 method works by disambiguating the current word's key-sequence to match the most likely word from a lexicon. Thus, often only one keypress is necessary per letter, but sometimes the wrong word is proposed and the user must select the desired word from a list. All sessions had subjects enter 12 test phrases.

In Parts 1 and 2, the order of methods was counterbalanced among subjects and across sessions. Also, no session was separated by less than 4 hours or by more than 2 days.

Part 3 of the study was run concurrently with the last 5 sessions of Part 1 (sessions 11-15). For this part, subjects entered 2 practice phrases followed by 8 test phrases using the back-mounted isometric joystick (Figures 1b, 4a).

Similarly, Part 4 of the study was run concurrently with the last 5 sessions of Part 2 (sessions 26-30). Using the same practice and test amounts as Part 3, subjects used the back-mounted joystick with word-level stroking (Figure 3b).

Finally, Part 5 of the study occurred at the very end, where subjects entered 5 test phrases with character-level EdgeWrite and Multitap while holding the phone beneath the edge of the table. Using the desktop PC, they were able to see the character they produced, but at no time could they see their hand, the device, or its screen. This simulated an eyes-free situation where feedback would be provided via some other means. Thus, subjects were "input blind" but not "output blind."

In all five parts, subjects transcribed a presented phrase in an "unconstrained" fashion [23, 30]. That is, they were free to enter the presented string using any combination of letters, spaces, and backspaces. Subjects were encouraged to proceed "quickly and accurately" [23] and to correct errors as they went.

Design and Analysis

Parts 1 and 2 of the study were analyzed separately as 2×15 within-subjects factorial designs with factors for *method* and *session*. These data were analyzed using a repeated measures ANOVA. Parts 3 and 4, which used the back-mounted joystick, were compared to the front joystick over the same sessions. Part 5 is analyzed using a single-factor within-subjects ANOVA. Subjective measures in the form of 5-point Likert scale responses are compared with non-parametric Wilcoxon signed-rank tests, although such strict tests present a challenge for detecting significance.

The assignment of methods to subjects was fully counterbalanced across subjects and across sessions. For each part of the study, we conducted tests for order effects, but none were significant, indicating adequate counterbalancing.

Measures

Results are reported for test phrases only. Dependent measures for a subject were mean session speeds (WPM) and uncorrected error rates (%) [23]. Corrected error rates were not analyzed because they are subsumed in speed, and because they are not appropriate for T9, where an entire word is first composed before it is "committed" all at once. Thus, our analyses consider how accurate the final text is, and how much time it took to enter it.

Part 1 • Character-level EdgeWrite *vs.* Multitap

Speed

Over all sessions (1-15), subjects wrote at 7.74 WPM (σ=1.72) with character-level EdgeWrite and 8.38 WPM (1.05) with Multitap (Figure 6a). However, this difference was not significant ($F_{1,5}$=2.51, p=.17). It is not surprising that the mean for Multitap was higher overall, since gestures had to be learned in EdgeWrite. If we exclude just the first two EdgeWrite sessions, EdgeWrite's mean climbs to 8.18 WPM (1.35). If we consider only the last three sessions (13-15), subjects wrote at 9.88 WPM (0.13) with EdgeWrite compared to 9.29 WPM (0.18) with Multitap. This suggests the presence of a crossover point, which is supported by a significant effect of *session* ($F_{14,70}$=91.98, p<.0001) and *session*method* ($F_{14,70}$=9.35, p<.0001). The maximum session speeds were 10.03 WPM for EdgeWrite (session 15) and 9.45 WPM for Multitap (session 13).

Error Rate

Both text entry methods had low uncorrected errors: 1.34% (σ=0.43%) for EdgeWrite and 0.52% (0.24%) for Multitap. As with speed, this difference was not significant ($F_{1,5}$=5.53, p=.07), although the trend was in favor of Multitap. But again we see that much of the difference occurred in the first two sessions when subjects were learning EdgeWrite and errors were highest, about 2%. Despite this change, there was no significant effect of *session* ($F_{14,70}$=0.98, p=.48) or of *session*method* ($F_{14,70}$=0.97, p=.50) on uncorrected errors.

Subjective Results

As measured by 5-point Likert scales filled out at the end of the study (Figure 6b), subjects felt that compared to Multitap, EdgeWrite was significantly faster (z=10.5, p<.05), marginally more enjoyable (z=9.0, p=.09), and marginally better liked (z=5.0, p=.12). However, they felt Multitap was marginally easier to learn (z=-7.5, p=.06). No differences were found for ease of use, perceived accuracy, or comfort.

Subjects indicated that they would unanimously select EdgeWrite over Multitap to enter "a few sentences," "a few paragraphs," and "a few pages of text." Overall, 5 of 6 subjects indicated they liked EdgeWrite more than Multitap. Some subjects also made written comments: "EdgeWrite is just like writing with a pen; very enjoyable," and "It is very intuitive since it very much corresponds to the way I write on paper." Of Multitap, subjects wrote: "This is too slow and boring," and "For any word with an 's' we have to tap four times, which is very tiring."

Part 2 • Word-level EdgeWrite *vs.* T9

Speed

Over all sessions (16-30), subjects wrote at 13.65 WPM (σ=1.68) with word-level EdgeWrite and 14.63 WPM (1.09) with T9 (Figure 6c). However, this difference was not significant ($F_{1,5}$=1.06, p=.35). As in Part 1, however, the early sessions required more learning for EdgeWrite than for T9. For example, if we exclude the first five sessions, EdgeWrite's speed increases to 14.57 WPM (0.92). As this suggests, subjects' speeds improved significantly over sessions ($F_{14,70}$=17.53, p<.0001), and marginally at different rates with each method ($F_{14,70}$=1.77, p=.06). We can see in Figure 6c that EdgeWrite's mean speeds seem to catch T9's

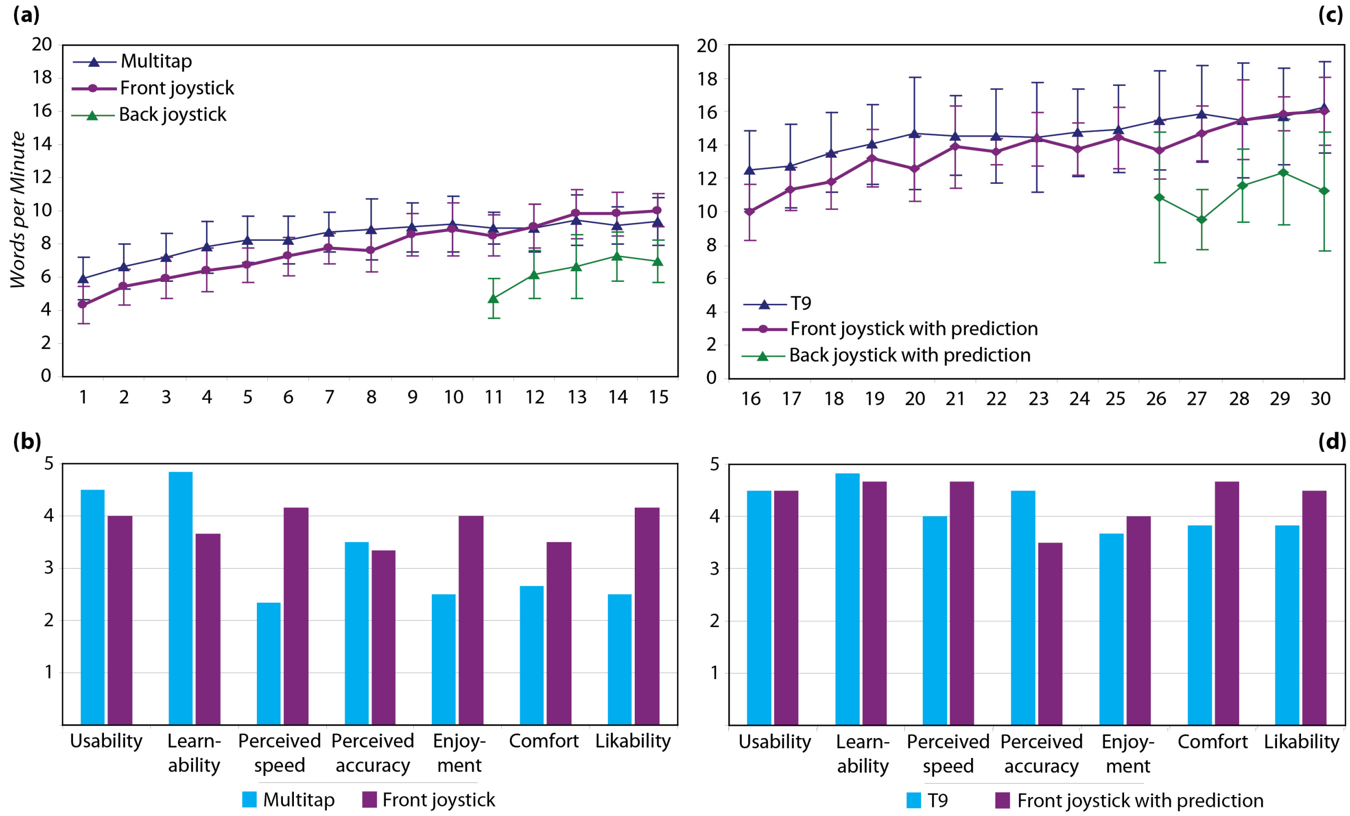

Figure 6. (a) Speeds of character-level EdgeWrite, Multitap, and the character-level back-mounted joystick. **(b)** Likert-scale measures for character-level EdgeWrite and Multitap. **(c)** Speeds of word-level EdgeWrite, T9, and the word-level back-mounted joystick. **(d)** Likert-scale measures for word-level EdgeWrite and T9. For WPM plots, error bars represent ±1 stdev.

speeds[2] in session 28. The maximum mean session speeds were 16.03 WPM with EdgeWrite and 16.25 WPM with T9, both in session 30.

Error Rate

In going from character-level entry in Part 1 to word-level entry in Part 2, both text entry methods left fewer uncorrected errors. Error rates were 0.35% (σ=0.32%) for EdgeWrite and 0.28% (0.16%) for T9. This difference was not significant ($F_{1,5}$=0.32, p=.60). As in Part 1, this difference occurred in the early sessions as subjects were learning word-level EdgeWrite. If we consider sessions 18-30, the means become 0.25% (0.21%) for EdgeWrite and 0.29% (0.17%) for T9. This creates a significant effect of *session* ($F_{14,70}$=2.28, p<.05) and *session*method* ($F_{14,70}$=2.07, p<.05), which confirms that subjects grew more accurate with the word-level techniques over time.

Subjective Results

Additional 5-point Likert scales were used to capture subjective results for the word-level techniques at the end of the study (Figure 6d). Subjects felt that compared to T9, EdgeWrite was marginally more comfortable (*z*=5.0, p=.12). No other differences were detected, although on average, EdgeWrite was perceived as faster, more enjoyable, more liked, but also more error prone.

Five of 6 subjects said they would select word-level EdgeWrite over T9 to enter "a few sentences" and "a few pages of text." All six preferred EdgeWrite for entering "a few paragraphs." Overall, 5 of 6 subjects indicated they liked word-level EdgeWrite more than T9. Subjects comments about word-level EdgeWrite included: "Very convenient for long words, but it is sometimes easy to make an error," and "Slightly frustrating when we enter a word completion by mistake," and "At times, one has to erase a word completion and do it again, wasting time." Of T9, subjects only wrote that it was "substantially better than [Multitap] because we do not have to tap repeatedly."

Part 3 • Character-level Back-Mounted Joystick

Speed

As a feasibility exploration, the back-mounted joystick was used for character-level EdgeWrite at the end of each session for the last 5 sessions in Part 1 (sessions 11-15). Thus, we can compare its performance to the front joystick for the same sessions[3] (Figure 6a). Overall means for these 5 ses-

[2] Note that out-of-vocabulary words were not automatically added to T9 during the study, but most tested words were in-vocabulary.

[3] Clearly, the fact that subjects had used the front joystick for 10 sessions prior to using the back-mounted joystick makes such comparisons unfair, but for simply exploring the feasibility of the back joystick in the first place, such comparisons will suffice.

sions were 6.34 WPM (σ=0.99) for the back joystick and 9.44 WPM (0.64) for the front joystick. Thus, the back joystick achieved 67.2% of the front joystick's mean speed over the same sessions. This difference was significant in favor of the front joystick ($F_{1,5}$=461.78, p<.0001). The peak back joystick speed was 7.27 WPM in session 14. Also, there were significant effects of *session* ($F_{4,20}$=25.72, p<.0001) and *session*method* ($F_{4,20}$=3.16, p<.05). Compared to Multitap's speed for these sessions (9.16 WPM, σ=0.23), the back joystick was slower ($F_{1,5}$=23.72, p<.01).

Error Rate

Over all 5 sessions (11-15), uncorrected error rates were 4.38% (σ=2.24%) for the back joystick and 1.45% (0.18%) for the front joystick. This difference was not quite significant ($F_{1,5}$=5.78, p=.06), which is somewhat surprising given the unique posture involved in using the back isometric joystick (Figure 4a). There were no significant effects of *session* ($F_{4,20}$=2.72, p=.06) or *session*method* ($F_{4,20}$=2.70, p=.06), although both were marginal since the first session with the back joystick had more errors—8.28%, compared to just 3.41% on average thereafter. Compared to Multitap's uncorrected error rate for these sessions (0.51%, σ=0.26%), the back joystick was less accurate ($F_{1,5}$=11.98, p<.05).

Subjective Results

Subjects preferred the front isometric joystick to the back-mounted joystick for ease of use (*z*=10.5, p<.05), perceived errors (*z*=10.5, p<.05), enjoyment (*z*=10.5, p<.05), comfort (*z*=10.5, p<.05), likeability (*z*=10.5, p<.05), and marginally for perceived speed (*z*=7.5, p=.06). No difference was found for ease of learning. These results are not surprising given subjects' longer practice with the front joystick, its familiarity, and its more customary position on the front of the device. The comments provided by subjects on the back joystick were: "frustrating" and "takes time to get used to."

Subjects preferred Multitap to the back joystick only for ease of use (*z*=10.5, p<.05), ease of learning (*z*=10.5, p<.05), and perceived errors (*z*=10.5, p<.05). Other responses—perceived speed, enjoyment, comfort, and likeability—were not different. Interestingly, subjects felt that the back joystick entry speed was *faster* on average than Multitap's speed (2.5 *vs*. 2.3 Likert response).

Part 4 • Word-level Back-Mounted Joystick

Speed

Part 3 was a further exploration of the back-mounted joystick, this time with word-level stroking enabled (Figure 3b). Trials were held during the last 5 sessions of Part 2 (sessions 26-30). As before, we can examine this design's performance to the front joystick for the same sessions (Figure 6c). Overall means for these 5 sessions were 11.11 WPM (σ=1.05) for the back joystick and 15.15 WPM (0.97) for the front joystick. This means that the back joystick achieved 73.3% of the front joystick's speed for these sessions. This difference was significant in favor of the front joystick ($F_{1,5}$=32.24, p<.01). The peak back joystick speed was 12.37 WPM in session 29. As in Part 3, there was also a significant main effect of *session* ($F_{4,20}$=4.11, p<.05), but this time, there was not a significant *session*method* interaction ($F_{4,20}$=1.27, p=.32). Compared to T9's speed for these sessions (15.75 WPM, σ=0.32), the back joystick was slower ($F_{1,5}$=12.84, p<.05).

Error Rate

Over all 5 sessions (26-30), uncorrected error rates were 0.20% (σ=0.18%) for the back joystick and 0.18% (0.23%) for the front joystick. This difference was not significant ($F_{1,5}$=0.02, p=.90). Due to a violation of sphericity, the Greenhouse-Geisser correction was used to interpret the main effect of *session*, which was not significant ($F_{2.1,10.6}$=3.68, p=.06). This correction was not necessary for the *session*method* interaction, which was significant ($F_{4,20}$=3.27, p<.05). Compared to T9's uncorrected error rate for these sessions (0.26%, σ=0.16%), the back joystick was not detectably different ($F_{1,5}$=0.16, p=.71).

Subjective Results

With word prediction enabled, subjects significantly preferred the front joystick to the back joystick for comfort (*z*=10.5, p<.05) and likeability (*z*=10.5, p<.05). They marginally preferred the front joystick for ease of use (*z*=7.5, p=.06), ease of learning (*z*=7.5, p=.06), perceived speed (*z*=7.5, p=.06), and enjoyment (*z*=9.0, p=.09). Only concerning perceived errors was there no detectable difference, which corresponds to the lack of a significant difference in actual error rates described above. Some comments by subjects were: "Easier and faster than before, but uncomfortable position for holding the device," and "better than the earlier one without word prediction."

Subjects significantly preferred T9 to the back isometric joystick for ease of use (*z*=10.5, p<.05), perceived speed (*z*=10.5, p<.05), perceived errors (*z*=10.5, p<.05), and comfort (*z*=10.5, p<.05). For ease of learning (*z*=7.5, p=.06) and likeability (*z*=7.5, p=.06), differences were marginal. Subjects did not exhibit a significant preference concerning enjoyment.

Part 5 • Device Held Under the Table ("Input Blind")

In Part 5, subjects entered 5 phrases with character-level EdgeWrite using the front isometric joystick and Multitap. Subjects did not use T9 simply because it would be impossible to do so without seeing the device. For this part of the study, the device was held under the edge of the table and out of sight. Under these circumstances, subjects might be called "input blind," but not "output blind," since they could see their results on a desktop PC screen. The goal was to compare performance when visual attention on the device was compromised. The Multitap phone had sufficient keypad tactility to support eyes-free use [21].

Speed

Speeds for entry beneath the table were 8.09 WPM (σ=1.04) for character-level EdgeWrite and 3.09 WPM

(0.58) for Multitap (Figure 7a). A single-factor within-subjects ANOVA for *method* yields a significant result in favor of EdgeWrite ($F_{1,4}$=126.19, p<.001). Note that EdgeWrite's 8.09 WPM is 80.7% of the maximum EdgeWrite speed achieved on the front joystick "above the table" (session 15). In contrast, Multitap's 3.09 WPM is only 32.6% of its peak speed (session 13).

Error Rate

Uncorrected error rates were 2.96% (σ=2.70%) for EdgeWrite and 1.58% (2.17%) for Multitap (Figure 7b). This difference was not significant ($F_{1,4}$=0.56, p=.49).

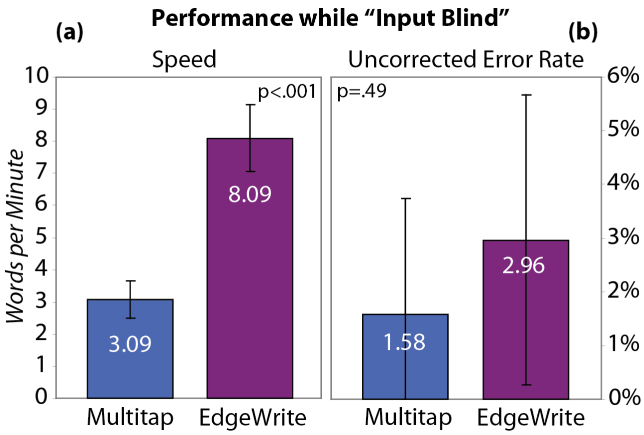

Figure 7. (a) Speeds of Multitap and character-level EdgeWrite while the device was held out of sight. EdgeWrite was significantly faster. **(b)** Uncorrected error rates showed no significant difference.

DISCUSSION

The major findings of this multifaceted study comparing our isometric joystick techniques to predominant keypad-based mobile phone text entry methods are:

- Gestural text entry on an isometric joystick was, in fact, feasible. Moreover, it seemed to be quickly learnable when using the EdgeWrite alphabet.
- Although no experiment can prove "equality," we can confidently say that the character-level version of EdgeWrite was "highly competitive" with Multitap. With six subjects and 15 sessions, there was ample power for one method to significantly outperform the other, but speed and accuracy statistical tests were not significant.
- Similarly, word-level EdgeWrite was very similar to T9 in speed and accuracy. Thus, including stroke-based word prediction in EdgeWrite gave rise to a method that can be said to be "highly competitive" with T9.
- Stroke-based word prediction is a viable means of improving gestural text entry speeds and error rates. This is notable since some previous word prediction systems have actually slowed users down [5, 22].
- Subjects showed a preference for both versions of EdgeWrite over their keypad-based counterparts. This shows subjects' willingness to endorse a gestural method despite the inevitable learning involved at the outset.
- Although slower, more error prone, and less preferred, using a back-mounted joystick *is* feasible and allows one to enter text at about 70% the front joystick's speed without a significant difference in uncorrected errors.
- When used without visual attention, our gestural method is about 2.5 times faster than Multitap.
- The speed of our gestural method does not degrade as much when used without visual attention as Multitap's speed does. The eyes-free gestural version is 80.7% as fast as in normal use, while eyes-free Multitap is only 32.6% as fast as in normal use.

At the outset of this study, we were somewhat skeptical that "daily intermittent use" would result in sufficient practice for EdgeWrite to eventually be competitive with Multitap and T9. Certainly, selection-based keypad methods were easier to use than gestural methods at the outset of the study. But we were somewhat surprised to see how quickly subjects learned EdgeWrite, and how their EdgeWrite results remained competitive with the results for two commercialized techniques over the course of the study.

In light of these findings, subjective results take on greater importance, since no method is an obvious "win." Here we see that subjects preferred EdgeWrite, particularly over Multitap, which they regarded as "tedious."

Extrapolated learning curves for all four techniques are shown in Figure 8. These are fit according to a power law and allow us to speculate about performance in future sessions. The curves suggest that the EdgeWrite methods have more long-term potential than the keypad-based ones, and that the speed advantage begun in the final sessions of Parts 1 and 2 would continue. Of course, further experimentation would be required to verify these predictions.

Some of the results from this study are similar to those found in prior studies. Butts and Cockburn [1] found that subjects reached 7.2 WPM with Multitap in a single session. This is similar to our result for Multitap in the third session. In work by Pavlovych and Stuerzlinger [17], subjects reached 7.15 WPM with Multitap after three 20-minute sessions. Our subjects reached 9.45 WPM after 13 sessions (~2 hours). James and Reischel [8] found that novices entered text with Multitap at 7.98 WPM while experts were about the same. For T9, novice speeds were 9.09 WPM while expert speeds reached 20.4 WPM. Our subjects did not reach this expert T9 speed by the end of their sessions, nor does their learning curve suggest they would anytime soon. Lyons *et al.* [11] ran a study that used Multitap in a longitudinal setting. They found that over twenty 20-minute sessions, two-handed Multitap speeds ranged from 8.2 WPM in the first session to 19.8 WPM by session 20. Our subjects could not reach these speeds because they did not use the methods for very long in each session, and they operated the phone with just one hand.

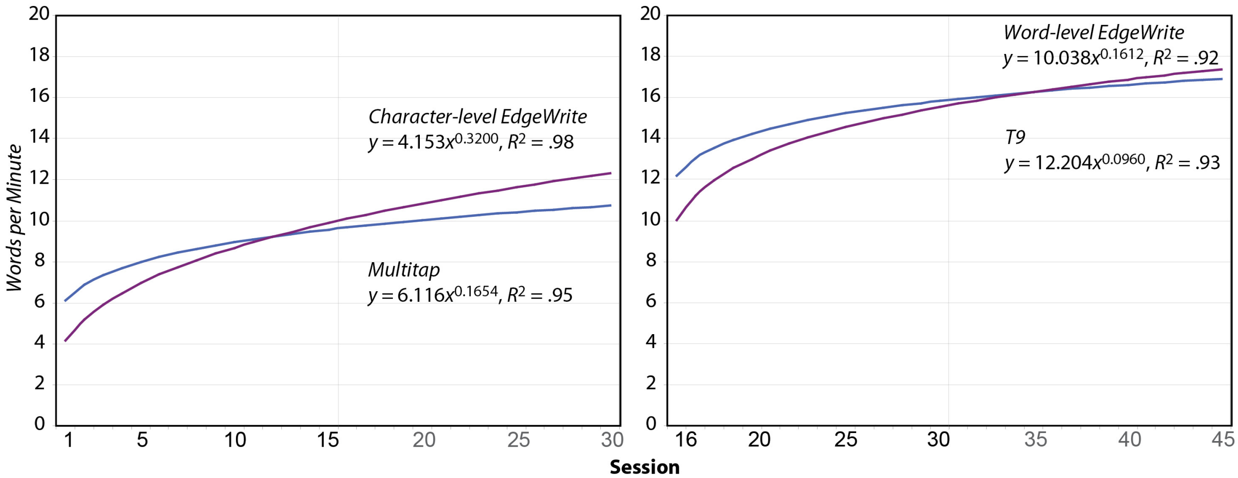

Figure 8. Fitted learning curves for the four main input techniques examined in this study. Curves are of the form $y = ax^b$.

Our isometric joystick design seemed to exhibit qualitative advantages such as the tiny physical footprint about the size of one button, the ability to perform fluid control tasks, its promise for gaming and multimedia, its ability to be more "felt" instead of "seen," and preference among subjects. Also, unlike some mobile methods where input and output are collocated on a display [14, 34], our method could function separately from where it is displayed, and the display can even be omitted. Of course, there are disadvantages to our joystick design as well. The main disadvantage is probably that regardless of how learnable a gestural method is, its gestures *still* must be learned. Device manufacturers may be reluctant to include gestural alternatives in lieu of keypad-based methods that, although tedious, require very little learning prior to use.

Compared to other versions of EdgeWrite [28], the *character-level* isometric joystick version (~10 WPM) seems slower than the stylus version (~24 WPM), the touchpad version (~19 WPM), the 4-key version (~17 WPM), and the version for game controllers (~15 WPM). However, it seems comparable to a version for power wheelchair joysticks and for trackballs (~13 WPM). Note that these character-level speeds reflect expert use as reported in [28] and should only be used for ballpark comparisons.

FUTURE WORK

Although our isometric joystick prototype fared competitively in the current study, there are still many ways in which it could be improved. Some of the subjects' comments indicated that word-level EdgeWrite had accuracy troubles when selecting words. Although erroneously completed words can be undone in EdgeWrite with a single backspace stroke, the design for word selection could be revisited in an effort to improve selection accuracy.

Numerous further studies would also be useful. Clearly, more sessions would help us assess what happens beyond sessions 15 and 30. We could evaluate our prototype against keypad-based techniques with subjects who are actually walking. We could also see how the techniques fare when they are not only "input blind," but "output blind" as well (i.e., subjects cannot see the results of their actions).

Finally, given the small size of the isometric joystick, other devices might benefit from having one of them. For example, a futuristic "pen" might have a joystick on its top used for pointing and text entry on remote displays, or a key-fob might use one to enter a few-letter password before unlocking a car to prevent signal spamming with a lost key.

CONCLUSION

In this paper, we investigated using an isometric joystick on both the front and back of a mobile phone for gestural text entry with the EdgeWrite alphabet. We compared our phone prototype to Multitap and T9, finding that our gestural text entry techniques were highly competitive and realistic alternatives to keypad-based techniques. Furthermore, subjects generally preferred our gestural designs to the tedious keypad-based systems, and our gestural system seems to have advantages for eyes-free use. This work has anticipated the future of highly capable mobile phones in an attempt to broaden their interaction vocabularies by using isometric joysticks to improve text entry.

Acknowledgements

The authors thank Brandon Rothrock, Joe Rutledge, and Ted Selker for their help. This work was supported in part by IBM, 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 or any other supporter.

REFERENCES

[1] Butts, L. and Cockburn, A. (2002) An evaluation of mobile phone text input methods. *Proc. 3rd Australasian User Interface Conf., Vol. 7 (AUIC '02)*. Sydney, Australia: Australian Computer Society, 55-59.

[2] Card, S.K., English, W.K. and Burr, B.J. (1978) Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys for text selection on a CRT. *Ergonomics 21* (8), 601-613.

[3] Chau, D.H., Wobbrock, J.O., Myers, B.A. and Rothrock, B. (2006) Integrating isometric joysticks into mobile phones for text entry. *Ext. Abs. CHI '06*. New York: ACM Press, 640-645.

[4] Coleman, M. (2001) Weegie. Available at http://weegie.sourceforge.net.

[5] Goodenough-Trepagnier, C., Rosen, M.J. and Galdieri, B. (1986) Word menu reduces communication rate. *Proc. RESNA '86*. Wash., D.C.: RESNA Press, 354-356.

[6] Hirotaka, N. (2003) Reassessing current cell phone designs: Using thumb input effectively. *Ext. Abs. CHI '03*. New York: ACM Press, 938-939.

[7] Isokoski, P. and Raisamo, R. (2004) Quikwriting as a multi-device text entry method. *Proc. NordiCHI '04*. New York: ACM Press, 105-108.

[8] James, C.L. and Reischel, K.M. (2001) Text input for mobile devices: Comparing model prediction to actual performance. *Proc. CHI '01*. New York: ACM Press, 365-371.

[9] Kanellos, M. (2006) Microsoft scientists pushing keyboard into the past. *CNET News.com*.

[10] Karlson, A.K., Bederson, B. and Contreras-Vidal, J.L. (2006) Studies in one-handed mobile design: Habit, desire and agility. *HCIL Technical Report 2006-02*, University of Maryland, College Park, Maryland.

[11] Lyons, K., Starner, T., Plaisted, D., Fusia, J., Lyons, A., Drew, A. and Looney, E.W. (2004) Twiddler typing: One-handed chording text entry for mobile phones. *Proc. CHI '04*. New York: ACM Press, 671-678.

[12] MacKenzie, I.S. and Soukoreff, R.W. (2003) Phrase sets for evaluating text entry techniques. *Ext. Abs. CHI '03*. New York: ACM Press, 754-755.

[13] Mithal, A.K. and Douglas, S.A. (1996) Differences in movement microstructure of the mouse and the finger-controlled isometric joystick. *Proc. CHI '96.* New York: ACM Press, 300-307.

[14] Nesbat, S.B. (2003) A system for fast, full-text entry for small electronic devices. *Proc. ICMI '03*. New York: ACM Press, 4-11.

[15] Nichols, J. and Myers, B.A. (2006) Controlling home and office appliances using Smartphones. *IEEE Pervasive Computing 5* (3), 60-67.

[16] Parsons, L.M. and Shimojo, S. (1987) Perceived spatial organization of cutaneous patterns on surfaces of the human body in various positions. *Journal of Experimental Psychology 13* (3), 488-504.

[17] Pavlovych, A. and Stuerzlinger, W. (2003) Less-Tap: A fast and easy-to-learn text input technique for phones. *Proc. Graphics Interface '03*. Waterloo, Ontario: Canadian Human-Computer Comm. Soc., 97-104.

[18] Rutledge, J.D. and Selker, T. (1990) Force-to-motion functions for pointing. *Proc. INTERACT '90*. Amsterdam: North-Holland, 701-706.

[19] Salem, C. and Zhai, S. (1997) An isometric tongue pointing device. *Proc. CHI '97*. New York: ACM Press, 538-539.

[20] Silfverberg, M., MacKenzie, I.S. and Kauppinen, T. (2001) An isometric joystick as a pointing device for handheld information terminals. *Proc. Graphics Interface '01*. Toronto, Ontario: Canadian Information Processing Society, 119-126.

[21] Silfverberg, M. (2003) Using mobile keypads with limited visual feedback: Implications to handheld and wearable devices. *Proc. Mobile HCI '03*. Berlin: Springer-Verlag, 76-90.

[22] Soede, M. and Foulds, R.A. (1986) Dilemma of prediction in communication aids. *Proc. RESNA '86*. Wash., D.C.: RESNA Press, 357-359.

[23] 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. CHI '03*. New York: ACM Press, 113-120.

[24] Ward, M. (2006) Cebit rings mobile phone changes. *BBC News*. London, United Kingdom: March 11, 2006.

[25] Wilson, A.D. and Agrawala, M. (2006) Text entry using a dual joystick game controller. *Proc. CHI '06*. New York: ACM Press, 475-478.

[26] 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. UIST '03*. New York: ACM Press, 61-70.

[27] Wobbrock, J.O., Myers, B.A. and Aung, H.H. (2004) Writing with a joystick: A comparison of date stamp, selection keyboard, and EdgeWrite. *Proc. Graphics Interface '04*. Waterloo, Ontario: Canadian Human-Computer Comm. Soc., 1-8.

[28] Wobbrock, J.O. and Myers, B.A. (2005) Gestural text entry on multiple devices. *Proc. ASSETS '05*. New York: ACM Press, 184-185.

[29] Wobbrock, J.O. and Myers, B.A. (2006) From letters to words: Efficient stroke-based word completion for trackball text entry. *Proc. ASSETS '06*. New York: ACM Press, 2-9.

[30] Wobbrock, J.O. and Myers, B.A. (2006) Analyzing the input stream for character-level errors in unconstrained text entry evaluations. *ACM Transactions on Computer-Human Interaction (TOCHI) 13* (4), 458-489.

[31] Wobbrock, J.O. and Myers, B.A. (2006) Trackball text entry for people with motor impairments. *Proc. CHI '06*. New York: ACM Press, 479-488.

[32] Zaborowski, P.S. (2004) ThumbTec: A new handheld input device. *Proc. Conf. on New Interfaces for Musical Expression (NIME '04)*. National University of Singapore, 112-115.

[33] Zhai, S., Smith, B.A. and Selker, T. (1997) Dual stream input for pointing and scrolling. *Ext. Abs. CHI '97*. New York: ACM Press, 305-306.

[34] Zhai, S. and Kristensson, P. (2003) Shorthand writing on stylus keyboard. *Proc. CHI '03*. New York: ACM Press, 97-104.