

Graffiti vs. Unistrokes: An Empirical Comparison

Steven J. Castellucci and I. Scott MacKenzie

Department of Computer Science and Engineering
York University
Toronto, Ontario M3J 1P3 Canada
{stevenc, mack}@cse.yorku.ca

ABSTRACT

Unistrokes and *Graffiti* are stylus-based text entry techniques. While *Unistrokes* is recognized in academia, *Graffiti* is commercially prevalent in PDAs. Though numerous studies have investigated the usability of *Graffiti*, none exists to compare its long-term performance with that of *Unistrokes*. This paper presents a longitudinal study comparing entry speed, correction rate, stroke duration, and preparation (i.e., inter-stroke) time of these two techniques. Over twenty fifteen-phrase sessions, performance increased from 4.0 wpm to 11.4 wpm for *Graffiti* and from 4.1 wpm to 15.8 wpm for *Unistrokes*. Correction rates were high for both techniques. However, rates for *Graffiti* remained relatively consistent at 26%, while those for *Unistrokes* decreased from 43% to 16%.

Author Keywords

Text entry, stylus, Graffiti, Unistrokes, empirical study.

ACM Classification Keywords

H.5.2 Information interfaces and presentation (e.g., HCI): User Interfaces---evaluation/methodology.

INTRODUCTION

Stylus-based entry techniques facilitate one-handed text input on portable systems, such as PDAs and tablet PCs. The user employs a pen-like stylus to “write” on a touch screen or digitizing tablet. The resulting “digital ink” can form gestures that are interpreted as text.

After introducing *Graffiti* and *Unistrokes*, we detail a user study to evaluate and compare them. We then present the results and elaborate on the findings.

Graffiti

Created and marketed by *Palm, Inc.* (www.palm.com), *Graffiti* (now *Graffiti2*) allows text entry with a stylus. A predominant feature of the *Graffiti* gesture alphabet

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI 2008, April 5–10, 2008, Florence, Italy.

Copyright 2008 ACM 978-1-60558-011-1/08/04...\$5.00.

(Figure 1, top) is that each stroke resembles its assigned Roman letter. This is intended to facilitate learning. Support for this was found in a previous study, where users demonstrated 97% accuracy after only five minutes of practice [6].

Other alphabets (e.g., *Jot*) employ a similar approach, but incorporate multiple, subtly different gestures for each Roman letter. This increases the chance that gestures match the user’s own handwriting, further facilitating proficiency. *Graffiti* is common in both commercial PDAs and in academic research [3].

Unistrokes

First introduced at the ACM SIGCHI conference in 1993, *Unistrokes* is a gesture alphabet for stylus-based text entry [2]. The term “unistrokes” characterizes all single-stroke gesture alphabets (*Graffiti* included). However, in this paper, “*Unistrokes*” refers specifically to the original gesture alphabet (Figure 1, bottom).

The single-stroke nature of each gesture allows entry without the user attending to the writing area [2]. Furthermore, the alphabet’s strokes are well distinguished in “sloppiness space” [2], allowing for accurate recognition of not-so-accurate input.

Unlike *Graffiti*, *Unistrokes* gestures bear little resemblance to Roman letters. However, each letter is assigned a short stroke, with frequent letters (e.g., E, A, T, I, R) associated with a straight line. *Unistrokes* is analogous to touch-typing with a keyboard, as practice will result in high-speed, “eyes-free” input [2].



Figure 1: The *Graffiti* (top) and *Unistrokes* (bottom) gesture alphabets.

METHOD

Participants

The ten paid participants (four males and six females) were students at the local university. They were recruited by posting flyers on campus. Ages ranged from 19 to 30 years ($mean = 23$; $SD = 3.07$). Of the ten, two were left-handed. Two used stylus-based devices at least once a week, three used them less frequently, and five had never previously used such devices. Nine frequently took hand-written lecture notes, while one favoured typing them. All possessed the dexterity to operate a stylus easily. None was familiar with either *Unistrokes* or *Graffiti*.

Apparatus

Figure 2 depicts the Java program used for gesture recognition and gathering text entry metrics. The topmost text area displayed the presented phrase, and the lower one the participant's transcribed text. The rectangle below is the stroke recognition area (SRA). The recognizer was borrowed from an earlier study [3]. Below the SRA is the enter button to terminate entry of a phrase. Above the SRA, reside the chart button and the backspace button. The absence of a backspace gesture in the *Unistrokes* paper [2] motivated the use of a backspace button.

The workstation for the experiment was a Pentium 4 530 (3 GHz) system with a *Wacom* PL-400 digitizing tablet, which integrates a 1024×768 LCD display. The workstation ran the text entry program using version 1.5.0 of the Java Runtime Environment. The program window was maximized to avoid extraneous onscreen stimuli. No additional applications were running. The study took place in a quiet office environment.

Procedure

At the beginning of each session, participants were given up to two minutes (five for the first session) to study a chart (similar to Figure 1) of the assigned gesture alphabet. By pressing the chart button, participants could view the chart in a popup modal dialog during the session. Since PDAs

come with gesture alphabet reference cards, we believe this aided external validity.

After being instructed to "enter the presented phrases as quickly and accurately as possible", participants used the stylus to enter text using gestures of the assigned technique. Gestures were entered in the SRA of the interface. The presented phrase remained visible to the participant throughout input. This aimed to eliminate errors due to spelling mistakes, and delays caused by forgetting a memorized phrase. Participants were allowed to rest between phrases. To encourage attention to the task, participants were required to correct all errors by using the backspace button and re-entering incorrect or mis-recognized gestures. Phrases with errors remaining were immediately repeated. Such erroneous phrase entries were not analyzed.

Design

The experiment was a 2×20 factorial design. A between-subjects factor, Input Technique, had two levels: *Graffiti* and *Unistrokes*. Use of a between-subjects factor eliminated any interference effects between the two techniques. A repeated measures factor, Session, represented twenty sessions of text entry. The dependent variables (and units) were Entry Speed (words per minute), Correction Rate (%), Chart Views (seconds per phrase), Stroke Duration (milliseconds), and Preparation Time (milliseconds).

The ten participants were randomly divided into two equal groups – one for each gesture alphabet. They made session appointments at their convenience of about fifteen minutes each. Sessions were separated by at least one hour but not more than three days. Each session involved fifteen phrases of text entry. Phrases were chosen randomly (without replacement) from a 500-phrase set [4]. The few instances of capital letters were converted to lowercase. The study lasted six weeks.

Entry time for each phrase was measured from the first pen-down event to the last pen-up event. It also included any time spent viewing the gesture alphabet chart. Using the accepted word length of five characters (including spaces) [8, p. 182], entry speed was calculated by dividing the phrase length by the entry time (in seconds), multiplying by sixty (i.e., seconds in a minute), and dividing by five (i.e., word length). Because participants were required to correct all errors, error rate was inherently zero percent. Instead, we calculated correction rate, defined as the number of backspace button presses per phrase divided by the length of the phrase.

RESULTS AND DISCUSSION

Performance

Figure 3 shows the results for Entry Speed. Learning effects were observed in both the *Graffiti* and the *Unistrokes* groups. During the first session, *Graffiti* users entered text at 4.0 wpm ($SD = 1.44$), while *Unistrokes* users entered at 4.1 wpm ($SD = 2.18$). By the twentieth session, *Graffiti*

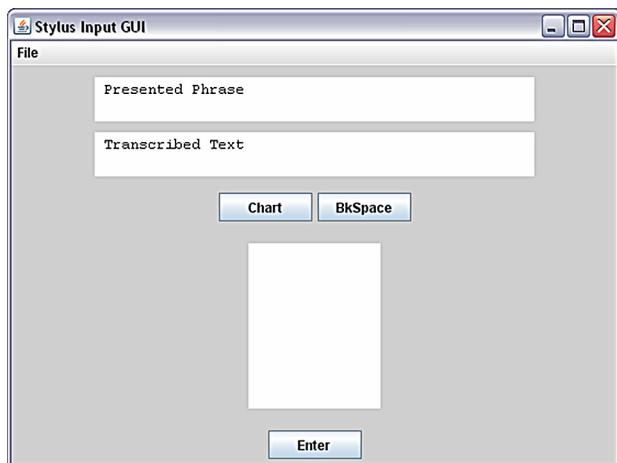


Figure 2: The interface used to gather text-entry metrics.

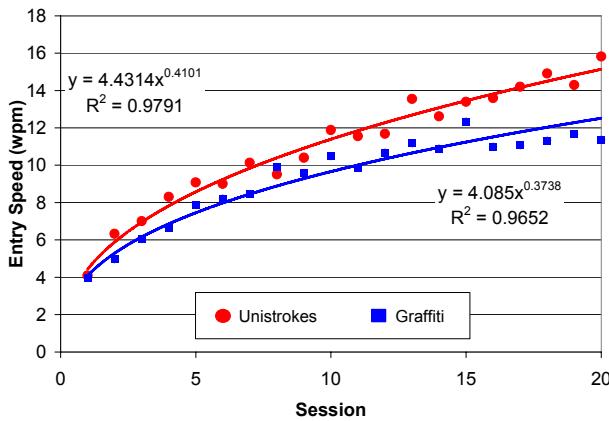


Figure 3: Entry Speed results.

users entered at an average of 11.4 wpm ($SD = 3.60$), while *Unistrokes* users entered at 15.8 wpm ($SD = 4.02$).

Graffiti's similarity with Roman letters suggests an advantage during initial sessions, whereas the simplicity of *Unistrokes* gestures lends itself to rapid input with practice. Though the Session \times Input Technique interaction effect was significant ($F_{19,152} = 2.26, p < .005$), the main effect of Input Technique on Entry Speed was not ($F_{1,8} = 2.05, p > .05$).

Correction Rate and Chart Views

Figure 4 illustrates the change in Correction Rate over the twenty sessions. Rates for *Graffiti* remained steady, averaging 26.2% ($SD = 2.6$). Those for *Unistrokes* decreased from 43.4% ($SD = 16.4$) for session one to 16.3% ($SD = 10.0$) for session twenty. Although these rates seem high, they can be explained as follows.

In addition to single backspace events in the logs, consecutive backspace events were also evident. These occurred because participants tended to view the writing area, as consistent with handwriting. Consequently, participants often missed errors. Once noticed, he or she repeatedly pressed the backspace button to perform the correction, deleting correct characters in the process.

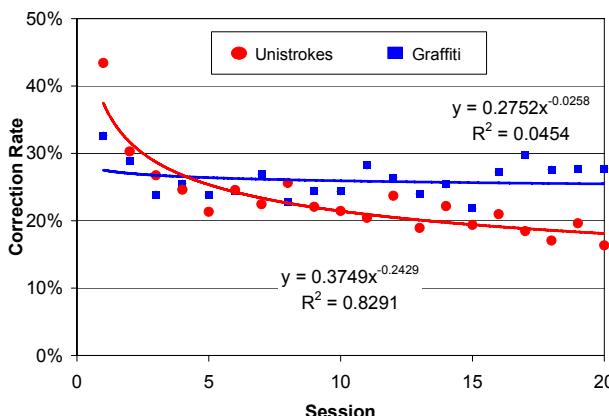


Figure 4: Correction Rate results.

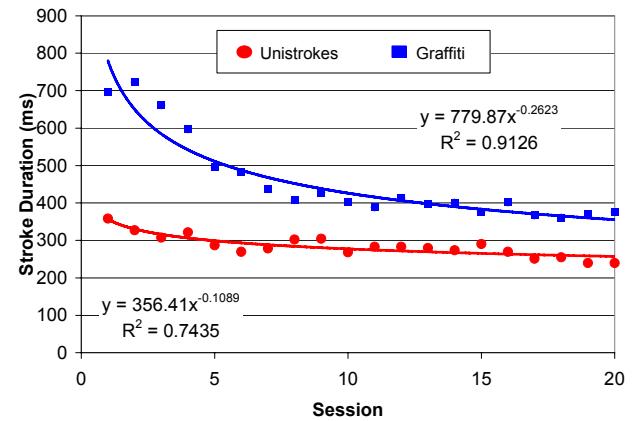


Figure 5: Stroke Duration results.

We also observed participants making repeated attempts to enter and correct a problem gesture. Instead of viewing the alphabet chart, participants favoured a guess-and-check approach. Viewing of the alphabet chart varied considerably. For the first session, *Graffiti* users spent an average of 4.0 seconds per phrase ($SD = 8.6$). For *Unistrokes* user, the average was much longer, at 12.5 seconds per phrase ($SD = 23.6$). For subsequent sessions, average chart viewing time for *Graffiti* users dropped to below one second per phrase, but the standard deviation remained high. The same was true for *Unistroke* users during the third and subsequent sessions.

Stroke Duration

Figure 5 displays the change in Stroke Duration (i.e., the time from pen-down to pen-up) over the twenty sessions. Due to their short and simple strokes, *Unistrokes* gestures were executed significantly faster than those of *Graffiti* ($F_{1,8} = 8.21, p < .05$).

Cao and Zhai devised a model of gesture composition by predicting the stroke duration of primitive components [1]. To evaluate it, they conducted an empirical study using *Graffiti* and *Unistrokes*. Table 1 presents a summary of their results. It also includes the stroke durations from this study, averaged over twenty sessions. Both studies yielded stroke durations much lower than the model's prediction. However, the empirical *Unistrokes-to-Graffiti* stroke duration ratios differ by only 0.32%! The discrepancy in the actual durations can be attributed to the additional practice afforded by this longitudinal study, compared to Cao and Zhai's small-scale [1, p. 1501] experiment.

Technique	This Study	Experiment Data[1]	Model Prediction[1]
Graffiti (A)	459	591	1125
Unistrokes (B)	284	365	622
Ratio (B/A)	0.620	0.618	0.553

Table 1. Stroke time data from this study and Cao & Zhai [1].

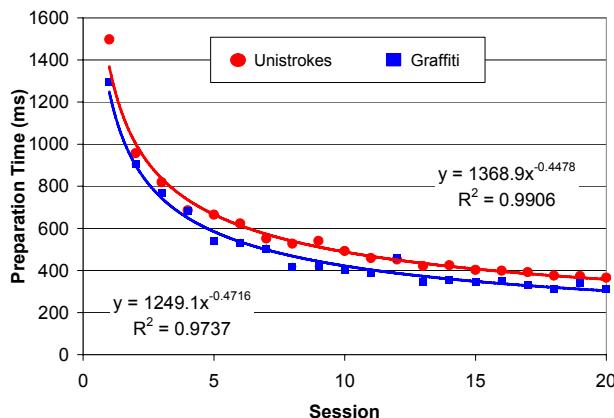


Figure 6: Preparation Time results.

Preparation Time

Figure 6 illustrates the change in Preparation Time (i.e., the time between strokes) over the twenty sessions. Again, the two curves exhibit obvious learning effects. However, while the improvement over the twenty sessions was statistically significant ($F_{19,152} = 54.20, p < .0001$), the main effect of Input Technique on Preparation Time was not ($F_{1,8} = 1.57, p > .05$).

The similarity of *Graffiti* to English suggests a significantly lower preparation time. Indeed, participants in the *Graffiti* group cited this feature as conducive to his or her performance. However, subtle differences between one's personal handwriting style and the *Graffiti* alphabet might diminish *Graffiti*'s mnemonic associations. For example, a common source of error in this study (and another [6]) was the addition of superfluous down-strokes in completing the letters G and U. While other gesture alphabets associate subtly different gestures with a single letter, the single gesture per letter design of *Unistrokes* makes *Graffiti* a better candidate for comparison.

Another explanation for the similar preparation times is *Graffiti*'s similarity with uppercase letters. Accounting for letter frequency, only 33% of lowercase Roman letters resemble their corresponding *Graffiti* gesture [6]. As typical of text entry experiments [4, 5], the current study employed phrases with only lowercase letters. Therefore, the presentation of such letters might have interfered with the strokes' mnemonic association with uppercase letters. This Stroop-like effect [7] would result in increased preparation time with the *Graffiti* alphabet. The *Unistrokes* alphabet lacks resemblance to Roman letters, and therefore is not susceptible to such an effect. Instead, its results can best be explained by participants' inexperience with the technique.

CONCLUSION

Over twenty fifteen-phrase sessions, text entry speed in the *Graffiti* group increased from 4.0 wpm to 11.4 wpm. During the same time, text entry speed in the *Unistrokes* group increased from 4.1 wpm to 15.8 wpm. However, an analysis of variance yielded a lack of statistical difference in entry speed between the two techniques. Participants often performed unnecessary deletions, resulting in high correction rates. In addition, the duration of gesture chart views decreased quickly, but varied widely between participants. Inter-stroke time between the two groups was similar, but the significant difference in stroke duration favoured *Unistrokes*.

The *Graffiti* alphabet's recognisability endears itself to novice users. However, this study shows that investing the same time learning *Unistrokes* can result in significantly faster stroke time and higher text entry speed.

ACKNOWLEDGMENTS

We wish to thank the CHI reviewers for their constructive feedback. This research was funded by the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

1. Cao, X. and Zhai, S. Modeling human performance of pen stroke gestures. In *Proc. CHI '07*, ACM Press, (2007), 1495-1504.
2. Goldberg, D. and Richardson, C. Touch-typing with a stylus. In *Proc. Interact '93 and CHI '93*, ACM Press, (1993), 80-87.
3. MacKenzie, I. S., Chen, J. and Oniszczak, A. Unipad: Single stroke text entry with language-based acceleration. In *Proc. NordiCHI '06*, ACM Press (2006), 78-85.
4. MacKenzie, I. S. and Soukoreff, R. W., Phrase sets for evaluating text entry techniques. *Ext. Abstracts CHI '03*, ACM Press (2003), 754-755.
5. MacKenzie, I. S. and Soukoreff, R. W. Text entry for mobile computing: Models and methods, theory and practice. In MacKenzie, I. S. ed., *Human-computer interaction* (2002), 147-198.
6. MacKenzie, I. S. and Zhang, S. X. The immediate usability of graffiti. In *Proc. GI '97*, Canadian Information Processing Society (1997), 129-137.
7. Stroop, J. R. Studies of interference in serial verbal reactions. *Journal of Experimental Psychology*, 18. 643-662.
8. Yamada, H. A historical study of typewriters and typing methods: From the position of planning Japanese parallels. *Journal of Information Processing*, 2 (4). 175-202.