A Comparison of Four Methods of Numeric Entry on Pen-based Computers

Craig McQueen[§], I. Scott MacKenzie[†], Blair Nonnecke[†], Stan Riddersma[†], Malcolm Meltz[¶]

[§]Computer Science Department
University of Toronto, Toronto
Ontario, M5S 1A4
jmcqueen@dgp.utoronto.ca,

[¶]Architel Systems Corp.
Toronto, Ontario, M8V 1A4
1-416-253-1920

[†]Dept. of Computing & Information Science
University of Guelph
Guelph, Ontario, N1G 2W1
mac@snowhite.cis.uoguelph.ca,
nonnecke@snowhite.cis.uoguelph.ca,
stan@snowhite.cis.uoguelph.ca

Abstract

Four numeric entry methods for pen-based computers were compared with respect to accuracy and speed. Sixteen subjects entered numbers on a digitizing display using four conditions: handwriting (with recognition software), keypad taping, pie pad, and moving pie menu. Keypad tapping was the most accurate (98.8%) and fastest (30.4 wpm) entry method. It was also most preferred by subjects. Handwriting was nearly as preferred as keypad tapping, even though it was substantially less accurate (89.6%) and slower (18.5 wpm). The pie menu conditions were least preferred, least accurate, and slowest. However, some subjects did achieve superior performance on the pie menu conditions; thus, with sufficient practice, pie menu entry may be a valid alternative to handwriting for numbers. These results call into question the presumed superiority of handwriting as the preferred entry method on pen-based computers.

La vitesse et la précision de quatre méthodes d'entrée de données sur un ordinateur à stylo ont été comparées. Seize sujets ont écrit des chiffres sur un écran digital de quatres manières différentes conditions: écriture manuscrite (avec l'aide d'un programme d'interprétation), usage des touches numériques, usage d'un pavé circulaire et usage d'un menu circulaire d'éplaçable. L'usage des touches numériques est la méthode la plus précise (98.8%) et la plus rapide (30.4 mpm). C'est aussi la méthode que les sujets de l'expérience ont préférée. La préference de l'écriture à la main est presqu'égale à celle des touches numériques mais la précision et la vitesse diminuent sensiblement (89.6% et 18.5 mpm). Les méthodes utilisant le pavé et le menu circulaires ont été moins préférées, moins précises et les plus lentes. Cependant, certains sujets ont obtenu des résultats supérieurs en utilisant le menu circulaire. Ainsi, avec suffisamment de pratique, le menu circulaire pourrait être une alternative à l'écriture manuscrite. Ces résultats mettent en question la supériorité de l'écriture manuscrite comme méthode d'entrée préférée pour l'ordinateurs à stylo.

Keywords: pen-based computing, gestures, pie menu, keypad, numeric entry,

Introduction

Pen-based computers have received much attention in the media recently, primarily due to new technologies entering the marketplace. They are appearing as machines running the *PenWindows* or *PenPoint* operating systems, Personal Digital Assistants (PDA), and large "whiteboard" displays. The potential market for pen-based computers includes people who work intensively with information and who work away from a desk (e.g., field-service personnel, couriers, doctors).

It has been suggested that the ability of pen-based technology to recognize handwriting makes it revolutionary and will change the way people enter information into computers. However, alternate methods of entry and the evaluation of these methods are rarely considered. In fact, products have already appeared such as the Amstrad *Pen Pad* (Pountain, 1993) that offer no alternative to handwriting entry.

Situations that rely only on numeric entry pose a more simplified problem than full text entry, since fewer symbols are needed. Regardless of the method of entry, empirical evaluations are necessary to determine which method is optimal for numeric entry with a pen computer. This paper compares four methods of numeric entry for pen-based computers.

Handwriting Recognition

Handwriting has received the most attention as an obvious and preferred input method for pen devices. Recent research and development efforts have produced commercial recognizers that convert the strokes of a printed character to an ASCII value. There are numerous recognizer "engines" on the market; for example, Gibbs (1993) surveys 13 recognizers from seven different vendors. Recognizers are most effective with block-printed characters. Performance improves by exploiting context, dictionaries, constrained symbol sets, user profiles, and training. Constraining the symbol set is particularly effective if numeric entry is expected in the application since the symbol set is reduced to ten. If the symbol set includes uppercase and lowercase letters, punctuation, and editing gestures, it can easily exceed 100 symbols, thus complicating recognition.

Since a benefit in using a pen is the skill transfer from handwriting, the performance of an "ideal" recognizer should be transparent to the user. That is, a perfect recognizer accepts and interprets natural handwriting at a rate controlled by the user, and the accuracy of the recognizer is equivalent to the accuracy of a human interpreting the writing. However, as Halfhill (1993) notes, "it'll be a long time before handwriting recognizers are as good as pharmacists at interpreting anybody's sloppy scrawl" (p. 74).

Accuracy of recognizers is the key to their success. Of the 13 recognizers surveyed by Gibbs (1993), seven quoted untrained walk-up accuracy of 92% for character-level recognition. Two cited rates of 85% and 90%. The remaining four cited rates of 85-90% for word-level recognition assisted by a standard dictionary. Gibbs (1993) also notes: "there is no accepted standard for evaluating accuracy. Each vendor assesses their own accuracy as they please" (p. 31). If these accuracy figures are correct then the performance of the recognizers on numeric-only entry should be much better since the character set is constrained.

Keypad Tapping

An alternative entry method is to select a digit from an on-screen keypad with a tap of the pen. This method exists in current graphical user interfaces, such as the calculator included with Microsoft *Windows* and the Apple *Macintosh.*

Impartial empirical tests of typing speeds and error rates for pen-tapping have not been published, although proprietary data are available (e.g., Carr & Shafer, 1991). The closest input scheme is a touch screen keyboard with text entry using fingers. For touch entry, Gould, Greene, Boies, Meluson, and Rasamny (1990) reported typing rates of 12 wpm; Wilkund and Dumas (1987) found speeds of 14-18 wpm with error rates under 1%. Different sizes of touchscreen keyboards were tested by Sears, Revis, Swatski, Crittenden, and Shneiderman (1993). They found entry rates of 10-21 wpm with a small layout and 20-32 wpm with a full-size touchscreen keyboard. Since a numeric keypad has less buttons to chose from, the speed should be somewhat higher.

A disadvantage of pen-tapping is the lack of kinesthetic feedback and the inability to have a reference point (Wilkund & Dumas, 1987). Hence, visual contact with the keypad is necessary during entry. Many handwriting recognition systems accept input anywhere on the digitizing surface, with text automatically sent to the application's "insertion point". Depending on the type of interaction required, this difference may be critical in selecting the input method. Consider the possible application of a border-crossing guard using a pen computer to enter license plate numbers: If the entry method required on-screen eye fixation, performance would be severely degraded. Tapping on a visual keypad would be a poor choice in this instance.

Gestures, Pie Menus, and Clock Strokes

Additional entry methods considered herein are based upon a novel concept that attempts to overcome the previous two methods' disadvantages. Since keypad entry forces eye fixation on the screen and handwriting entry suffers from accuracy problems, the investigation of other entry methods is warranted.

One of the most alluring claims for pen-based computing is that natural gestures can form the core repertoire of interaction techniques. Numerous studies have shown the tremendous potential of the pen as a gestural input device (e.g., Buxton, 1986; Hardock, 1991; Kurtenbach & Buxton, 1991; Wolf & Morrel-Samuels, 1987). Although researchers hasten to point out that alphanumeric symbols are germane to pen-based input, we should acknowledge that these are culturally biased and are learned only with considerable practice.

Goldberg and Richardson (1993) presented a system called "unistrokes" for text entry. Their system consisted of symbols that map one-to-one to the regular alphabet. The symbols consisted of a single stroke to make them fast to write and easy to recognize. They reported an average writing rate of 2.8 letters per second (33.6 wpm); however, only three subjects were tested. Perhaps a similar yet simpler entry method would be effective for numeric entry.

Callahan, Hopkins, Weiser, and Shneiderman (1988) investigated pie menus for selection. They showed that an appropriate organization within a pie menu improves performance. Kurtenbach, Sellen and Buxton (1993) found that increasing the number of slices in the pie monotonically increased response time, except for pie menus that contained 12 items, similar to a “clock face”. One of the authors' suggestions was that the clock metaphor may have reduced visual search time, thus reducing overall response time.

Combining unistrokes and pie menus with a clock metaphor yields a new single-stroke numeric entry method - clock strokes. That is, the user enters a number by stroking from an arbitrary starting point toward where that number occurs on the face of a clock. This is illustrated in Figure 1. A "3" is entered as a stroke to the right, a "6" as a downward stroke, and so on. Note that 0 is at the 12 o'clock position and the 10 and 11 o'clock positions are not used. The advantages are single-stroke entry, a strong metaphor to minimize learning, and the scripting of strokes anywhere at any size.

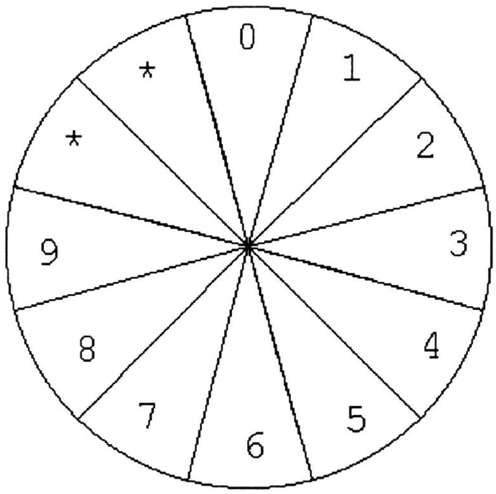

Figure 1. The clock metaphor for numeric entry

We investigated clock strokes two ways. One is to stroke where the digit is entered, similar to handwriting. We call this the *moving pie menu.* The other is to use a stroking pad where the user performs the strokes. Using this scheme, which we call a *pie pad,* all strokes are made on top of one another. The resulting digit is automatically sent to the application's insertion point. This method is what Goldberg and Richardson (1993) call "heads-up writing". The advantages of a pie pad are a reduced writing area, eyes-free operation, and less wrist fatigue.

In the following section, we describe an experiment that explores human performance in a pen-based numeric entry task using handwriting, a numeric keypad, a moving pie menu, and a pie pad. Performance is measured by the speed and accuracy of digit entry. It is expected that clock stroking is faster than handwriting after practice since less stroking is required and the strokes are simple. As well, it is expected that most errors for the clock strokes will occur on the "off-angle" digits (1, 2, 4, 5, 7, 8). The clock strokes should show the most improvement with practice since it is the method that subjects are least accustomed to.

Method

Subjects

Three female and 13 male volunteer subjects were used in the study. All were university students who used computers on a daily basis.

Apparatus

Software to run the experiment was developed in C using Microsoft's *Pen Windows.* Microsoft's handwriting recognition software (included with *Pen Windows)* was used and was configured to recognize the digit symbols only. *Pen Windows* time stamps the strokes according to when they where created thus discounting machine latency times.

Hardware for the experiment consisted of a 50 MHz PC-486 with a Wacom *PL-100V* tablet for pen entry. The *PL-100V* is both a digitizer for user entry and a 640 × 480 LCD screen. Using the combination of the tablet and host computer allowed the experiment to run without system lag due to the fast processor and allowed user entry to also appear on a regular VGA monitor. The monitor was tilted to prevent subjects from seeing it. Digits produced for user entry were generated using the random number generator library function provided by the C compiler.

The clock stroke conditions accepted the digits one to nine as strokes in the same direction as on the face a clock. Digit zero was assigned the 12 o'clock direction. If the user stroked in the 10 or 11 o'clock directions, it was recorded as not recognized. Each digit had a quantization space of 15° on either side of its ideal angle (0° for "0", 30° for "1", and so on).

Procedure

The task consisted of entering digits provided by the software using one of four conditions. The task is comparable to entering digits in a numeric field of a form. The conditions were (a) handwriting, (b) on-screen keypad, (c) moving pie, and (d) pie pad, as illustrated in Figure 2. Digits were presented randomly in groups of five. A group of five was called a sequence. Ten sequences made up a block.

Subjects performed all four conditions over two sessions of about 50 minutes each. The sessions took place over two days. Conditions were counterbalanced using a Latin square to minimize transfer effects.

Subjects were instructed to aim for both speed and accuracy when entering the digits. As well, subjects were told if a mistake was made they were to ignore it and continue with the sequence. The tablet was set flat on the table and subjects were told to rest their hand on the tablet so that tablet positioning would be consistent across subjects. In the case of the two clock stroke conditions, the idea was explained to subjects by showing them a picture of where the numbers were located in the pie. While stroking, the pie was not displayed.

The keypad and pie pad were windows that could be re-positioned (dragged). Subjects were allowed to relocate these to any location on the screen. This was particularly important for the left-handed subjects.

Execution of a condition consisted of a brief practice session of 10-15 sequences and then 20 minutes of recorded entry. To help motivate subjects, summary data for accuracy and speed were displayed at the end of each block. Typically, subjects completed 15 to 20 blocks for a total of 750 to 1000 digits within the allocated 20 minutes. A feedback click was produced upon the recording of a digit.

Two timing values were recorded for each digit, preparation time and scripting time. Preparation time is the time from the end of the previous digit to the start of the current digit. Scripting time is the amount of time that the pen is in contact with the tablet while forming the digit or pie stroke. Preparation time plus scripting time equaled the total entry time for a digit. Note that the keypad scripting time is virtually zero due to the brevity of the tap. As well, the timing value for the first digit in a sequence is meaningless because there is no start time to reference from. Thus, the data for the first digit in a sequence was not used for the summary statistics.

Results and Discussion

Condition Effects

Data were summarized for the first, middle, and last three blocks for each condition. The data entered in the analysis of variance were from the last three blocks only, to minimize learning effects. This data set contained 120 digits per subject for 16 subjects which totals 1920 digits for each condition.

There was a significant main effect for condition on entry time ($F_{3,45} = 59.3$, $p < .0001$) and error rate ($F_{3,45} = 22.1$, $p < .0001$). The mean values for each condition are illustrated in Figure 3. To facilitate comparisons with other studies, entry speed is shown in words per minute (wpm) in the figure. Note that the keypad is superior to the other methods for both accuracy and entry time. At 30.4 wpm, keypad tapping is comparable to Goldberg and Richardson's (1993) unistrokes or to Sears et al.'s (1993) full-size touchscreen keyboard.

Learning Effects

The improvement of speed and accuracy across sessions was investigated. Although there was no effect across blocks for accuracy ($F_{2,30} = 1.5$), there

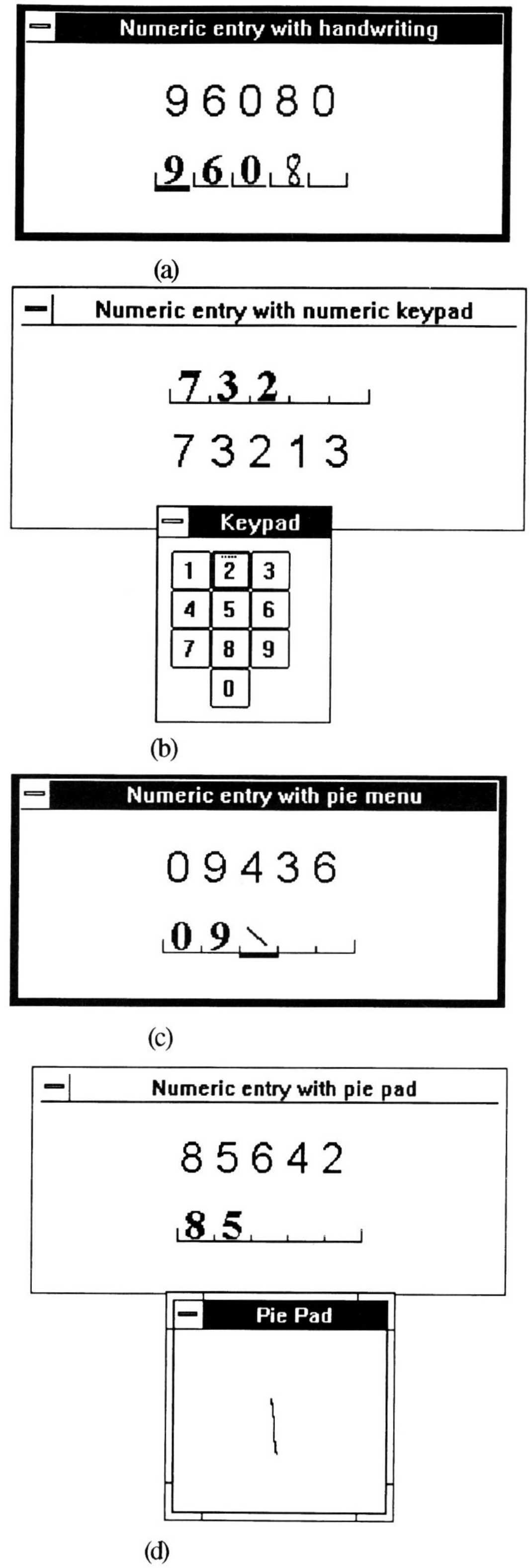

Figure 2. The four experimental conditions were (a) handwriting, (b) keypad, (c) moving pie menu, and (d) pie pad.

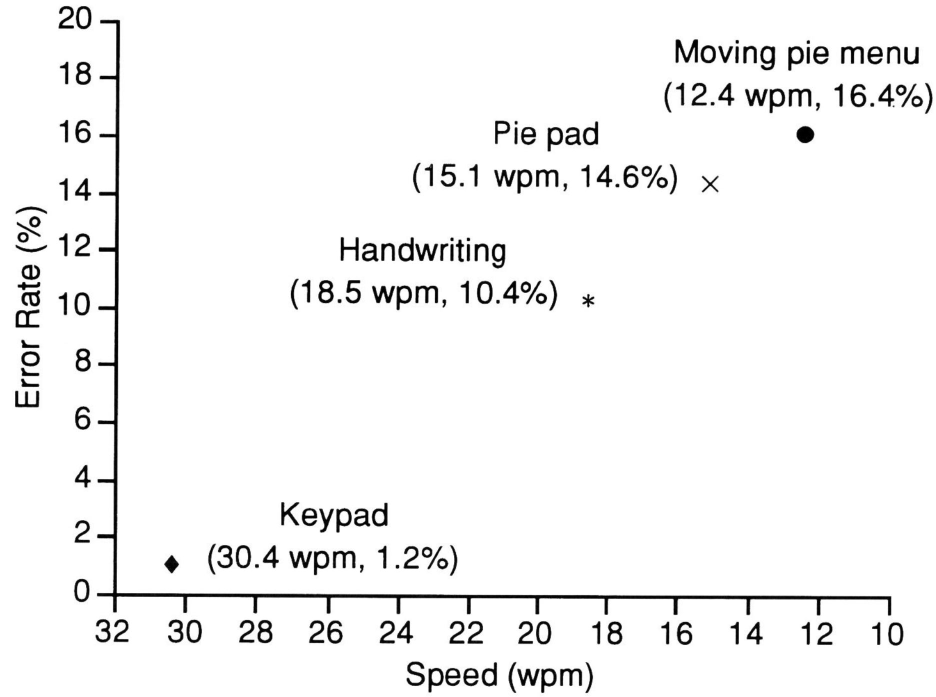

Figure 3. Comparison of the four conditions for error rates and entry time.

was significant main effect across blocks for entry time ($F_{2,30} = 14.2$, $p < .0001$). There was no noticeable improvement in accuracy for any of the conditions over the session. This is consistent with Bailey's (1989) observation that "in activities where performance is primarily automatic the proportion of errors will remain fairly constant, but the speed with which the activity is performed will increase with practice" (p.101).

The improvement in speed over the sessions relative to the initial performance was in the following order: keypad (14.1%), pie pad (11.3%), moving pie menu (6.2%), and handwriting (3.7%).

The keypad condition was the easiest condition to learn which is most likely the reason it had the largest decrease in entry time over the session. Handwriting is the most natural task, requiring minimal learning, and thus entry time reduction was not great. Entry time reduction was observed in the pie menu conditions. Greater performance with the pie menus would be observed over more practice sessions since these conditions are the most novel.

Note that the learning effects measured were only from one session of about 20 minutes for each condition. Prolonged use is merited to accurately assess learning time and retention with pie menus.

Performance by Digit

Since the entry technique varies across digits, analyses on a per digit basis are warranted. For handwriting, the digits 3 and 8 accounted for 43.4% of the errors. This is probably attributable to the recognizer, since the strokes for "3" and "8" are very similar. Indeed, our test of handwriting accuracy is more a test of the recognition ability of the particular product than a test of handwriting accuracy per se. Other recognizers would produce different results and merit investigation on their own.

With the keypad, 0 and 5 had the highest error rates: 1.93% and 2.41% respectively compared with a mean of 1.2%. Off-angle strokes accounted for the majority of errors in the clock stroke conditions. The lowest error rates were for the digits 0, 3, 6, and 9, as shown in Figure 4.

Preparation vs. Scripting Time

Although the total time to enter a digit was greater for the pie menu conditions than for handwriting, a comparison of preparation and scripting time reveals the potential of pie menus (see Table 1).

Preparation and total time were less for handwriting than for the pie conditions, while the stroking time during handwriting was greater by about 50%. Hence, the pie conditions would be faster than handwriting if preparation time was reduced. Preparation time is higher with the pie tasks because they require a conscious act. Handwriting, on the other hand, is a highly learned motor skill. In fact, subjects commented that the pie methods were fatiguing and demanded a lot of concentration. If practiced over a number of days, the pie menu tasks

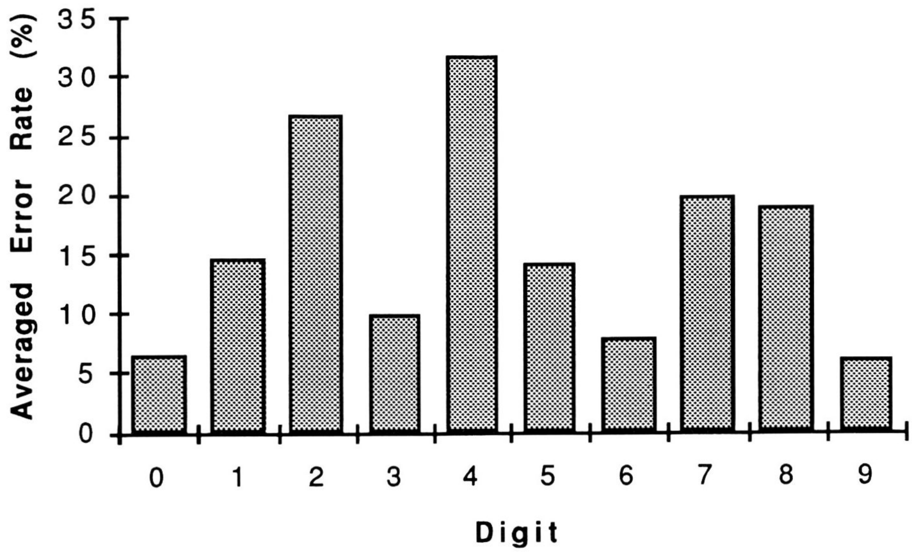

Figure 4. Error rates by digit for the pie conditions. The lowest error rates were for the on-axis digits: 0, 3, 6, and 9.

Table 1
Comparison of Preparation and Scripting Times (ms)

Condition	Preparation Time	Scripting Time	Total
Handwriting	334	315	649
Pie pad	569	226	794
Moving pie	763	207	970

would become more automatic with a reduced cognitive load. This, no doubt, would reduce preparation time.

Subject Differences

Superior performance for the pie pad condition was observed in two subjects. They achieved an accuracy of 99.2% and 95.0% for the pie pad verses 90.0% and 94.2% for handwriting respectively. Their times for the pie pad entry were 683 ms and 448 ms per digit verses 489 ms and 552 ms for handwriting. The first subject commented that he kept an image of a pie in his head. Improved performance seems to be achieved by concentrating less on the task. Bailey (1989) states that "a person cannot effectively plan and act at the same time" and "consciously analyzing performance while performing detracts from quickly accomplishing the activity. Movements are slowed" (p. 103). Thus, the key to effectively using pie strokes is a strategy that minimizes planning and leads to an automatic process. To be effective, pie strokes must become as automatic as handwriting.

A strategy that some subjects cited for the keypad condition was to memorize the five digits in the sequence prior to entering them. This allowed them to focus on entering the digits and resulted in decreased entry time. It also changed the task from a "copy" task to a "create" task since the subject was retrieving digits from memory rather than copying from another physical location.

During handwriting, some subjects improved their accuracy by adapting to the recognizer. For instance, if the recognizer was not accepting a subject's 4's, the subject would script them differently until the recognizer accepted the 4's. The subject then continued to script in this manner for the rest of the session. Subjects commented that they did not mind adapting to the recognizer and that it was sometimes unconscious.

Table 2
Condition Preference (frequency)

Condition	Rating[a] 5	4	3	2	1
Handwriting	1	11	4	0	0
Keypad	4	8	4	0	0
Moving pie	0	1	3	7	5
Pie pad	0	2	3	8	3

[a] 1 = least preferred, 5 = most preferred

Condition Preference

Subjects were asked to rate each condition on a one to five scale. The results are listed in Table 2. The keypad received the highest preference ratings, followed closely by handwriting. The pie pad was preferred over the moving pie.

Many of the subjects realized the keypad was a better entry method yet they preferred handwriting. They felt it was more natural, easy, and directly corresponded to the task. The implication of this is that if a new user is given the choice of handwriting or keypad entry, the user may chose handwriting even though it is less efficient.

Improving the Pie Menu

Often a subject's strokes were in the general direction of the number but the subject would curve the line either at the start or beginning. Since the digit assignment is generated by the angle formed from the start and end points, the digit could be quantized incorrectly if the trajectory curved into the wrong "slice". Other ways of determining the desired digit, such as using points 10% of the total distance from each end of the line, might improve the accuracy of the clock stroke methods.

Further research into pie divisions could improve the performance of the pie menu methods. For instance, the digits that have higher error rates could be assigned a larger slice of the pie. Investigation into using 10 instead of 12 sections could be done. This would eliminate the clock metaphor and reduce the method to a pie menu selection. It was found that the mean directions for the digits was not directly at the allocated angles (e.g., the mean stroking angle for 0 was at 5.6 degrees rather than 0 degrees). Perhaps using the actual mean angles as the centre of the slice would improve the accuracy. As well, digits with higher standard deviation values for stroking angle could be allocated larger slices of the pie.

Entry Differences

Even though the keypad had superior performancefor speed and accuracy it has two disadvantages: It consumes screen real estate, and users must look at the keypad while entering. Handwriting and the moving pie menu do not take up any screen real estate except for the place where the digits are entered. The user must still look where scripting is being performed, however. The pie pad has the same heads-up display advantage as unistrokes, but it requires screen real estate for scripting.

The error rate reported herein for handwriting (10.4%) was higher than the 8% walk-up error rate cited for Microsoft's recognizer in Gibbs' (1993) survey. The poor performance we observed would undoubtedly worsen with an unconstrained symbol set. This is a serious problem with handwriting recognition, and one that is played-out regularly in the popular press. Handwriting recognition has been called "impractical" (Eglowstein, 1993), "marginal" (Halfhill, 1993), and "disappointing" (Caruthers, 1993). Apple acknowledges walk-up error rates as high as 40-50% with their *MessagePad*, for example (Cassleman, 1993). Training a system to recognize an individual's handwriting is possible, but brings other problems, such as an inability to use any PDA, and the need to constantly back-up user profiles on a host system.

It is important to note that speed was measured independant of the recognizer and thus the handwriting times are just that, the amount of time it takes to script the digits. Technology will never improve handwriting speed but other novel entry methods might.

Conclusions

We found that keypad tapping is the best numeric entry method in terms of accuracy and speed of entry. The pie menu conditions on average did not perform as well as handwriting and keypad. However, some subjects achieved performance levels equal to or better than handwriting.

The effect of button size and inter-button space could be investigated further. This would allow optimization of keypad layout. Comparisons with other recognizers are necessary to determine if the accuracy of handwriting could be increased. For numeric entry using handwriting to be feasible, the performance of character recognizers must be improved beyond the current state.

An extended study of pie menu entry is required. Subjects would need to practice in many sessions for a greater period of time to completely learn the method. As well, optimization of the pie menu layout could be performed according to subject's stroking tendencies.

Exploration of text entry would explore similar issues as discussed in this paper except using a full alphanumeric symbol set, cursive script, or other features that will determine the success of pen-based input.

Acknowledgements

This research is supported by Architel Systems Corporation, the University Research Incentive Fund (URIF) of the Province of Ontario, and the Natural Sciences and Engineering Research Council (NSERC) of Canada. We gratefully acknowledge this support without which this research would not have been possible.

References

Bailey, R. W. (1989). Human performance engineering (2nd ed.). Englewood Cliffs, NJ: Prentice Hall.

Buxton, W. (1986). Chunking and phrasing and the design of human-computer dialogues. In H.-J. Kugler (Ed.), *Proceedings of the IFIP 10th World Computer Conference--Information Processing '86*, 475-480. Amsterdam: Elsevier Science.

Callahan, J., Hopkins, D., Weiser, M., & Shneiderman, B. (1988). An empirical comparison of pie vs. linear menus. *Proceedings of the CHI '88 Conference on Human Factors in Computing Systems*, 95-100. New York: ACM.

Carr, R., & Shafer, D. (1991). *The power of Penpoint*. Reading, MA: Addison-Wesley.

Caruthers, F. (1993, March). Pen input drives new class of personal computers. *Computer Design*, pp. 12-16. [OEM supplement]

Cassleman, G. (1993, August 16). Newton here but in short supply. *Computing Canada*. pp. 1, 4.

Eglowstein, H. (1993, July). Applying the power of the pen. *Byte*, pp. 132-140.

Gibbs, M. (1993, March/April). Handwriting recognition: A comprehensive comparison. *Pen*, pp. 31-35.

Goldberg, D., & Richardson, C. (1993). Touch-Typing with a Stylus. *Proceedings of the INTERCHI '93 Conference on Human Factors in Computing Systems*, 80-87. New York: ACM.

Gould, J. D., Greene, S. L., Boies, S. J., Meluson, A., & Rasamny, M (1990). Using a touchscreen for simple tasks. *Interacting with Computers, 1*, 59-74.

Halfhill, T. R. (1993, October). PDAs arrive but aren't quite here yet. *Byte*, pp. 66-86.

Hardock, G. (1991). Design issues for line-driven text editing/annotation systems. *Proceedings of Graphics Interface '91*, 77-84. Toronto: Canadian Information Processing Society.

Kurtenbach, G., & Buxton, B. (1991). GEdit: A testbed for editing by contiguous gestures. *SIGCHI Bulletin, 23*(2), 22-26.

Kurtenbach, G., Sellen A., & Buxton, B. (1993). An empirical evaluation of some articulatory and cognitive aspects of marking menus. *Human-Computer Interaction.*

Pountain, D. (1993, August). Amstrad's Modest PDA. *Byte*, 161-164.

Sears, A., Revis, D., Swatski, J., Crittenden, R., & Shneiderman, B. (1993). Investigating touchscreen typing: The effect of keyboard size on typing speed. *Behaviour & Information Technology, 12*, 17-22.

Wilkund, M. E., & Dumas, J. S. (1987). Optimizing a portable terminal, keyboard, for combined one-handed and two-handed use. *Proceedings of the 31st Annual Meeting of the Human Factors Society*, 585-589. Santa Monica, CA: Human Factors Society.

Wolf, C. G., & Morrel-Samuels, P. (1987). The use of hand-gestures for text-editing. *International Journal of Man-Machine Studies, 27*, 91-102.