Technical Report

IBM

Input Rates and User Preference for Three Small-Screen Input Methods: Standard Keyboard, Predictive Keyboard, and Handwriting

James R. Lewis

February 10, 1995

TR54.889

Unclassified

Boca Raton, Florida

Unclassified

Input Rates and User Preference for Three Small-Screen Input Methods: Standard Keyboard, Predictive Keyboard, and Handwriting

James R. Lewis

TR54.889

ABSTRACT

As touchscreen interfaces become smaller, emulating keyboard input becomes more difficult. One approach for solving this problem is to provide a small on-screen version of a standard keyboard. Another solution to keyboard emulation is a predictive keyboard -- one that only displays a subset of the full set of keyboard keys, with predictions based on tables of letter frequencies (usually based on letter triplets, or trigraph frequencies). Handwriting recognition is a third method for dealing with this problem. Each of these approaches has its benefits and drawbacks. The standard keyboard shows all the letter keys simultaneously, so the keys are always in the same place and the user can draw upon prior knowledge of key layout. However, if the device is small, this approach makes the user type with fairly small keys. The predictive keyboard's keys can be larger, but the user must always scan the keys to see which are available and must touch other keys to bring up letters, numbers, or punctuation not currently displayed. Current handwriting recognizers cannot recognize handwriting perfectly (but might be able to in the future). Previous research has shown that users prefer an on-screen keyboard to handwriting recognition when handwriting recognition is imperfect, so in this study the device only captured the users' handwriting. The experiment told the participants to assume perfect handwriting recognition.

The purpose of this study was to compare these approaches (standard keyboard, predictive keyboard, and perfect handwriting recognition) considering user input rates and preferences. With counterbalanced orders of presentation and stimuli, six participants used all three methods to enter both normal text and addresses. Participants indicated (with both rankings and ratings) that they preferred to use the standard keyboard. The average input rate for handwriting was fastest, but also much more variable than the other methods. Despite its speed, participants generally found it difficult to write comfortably and legibly on the small (35 x 115 mm) display. The input rate for the standard keyboard was more than twice the input rate for the predictive keyboard. These results suggest that, for small devices, neither handwriting recognition nor predictive keyboards can effectively replace the standard keyboard layout. Until developers can invent something better, the standard keyboard layout should be the default input method for small devices. Even with perfect handwriting recognition, users will prefer to tap on a small standard keyboard unless the device's hardware design allows comfortable handwriting input.

Introduction

As touchscreen interfaces become smaller, emulating keyboard input becomes more difficult. Providing a method for effective data input for small, handheld devices is one of the major usability problems facing personal digital assistants (PDAs) and personal communicators (PCs) (Lewis, 1994). With the introduction of Newton** and Simon**, there are three methods currently available: 1) an on-screen keyboard with a standard layout, 2) an on-screen predictive keyboard, and 3) handwriting recognition. Each of these approaches has its benefits and drawbacks.

The most obvious and easiest solution is an on-screen version of a standard (QWERTY) keyboard layout. In 1991, Sears reported that users could type an average of 25 words-per-minute (wpm) on touchscreen keyboards with keys 22.7 mm square, and could type an average of 20 wpm using a keyboard only 70 mm wide (implying keys about 7 mm square). As the size of the keys decreases, typing speed decreases (Sears, 1991). A number of studies have shown that the standard layout is superior to alphabetic layouts, regardless of user experience with standard layouts or size of keyboard (Francas, Brown, and Goodman, 1983; Norman and Fisher, 1982), so any full-keyboard layout should use the standard layout.

Another solution to on-screen keyboard emulation is a predictive keyboard -- one that only displays a most-likely subset of the full set of keyboard keys, with predictions based on tables of letter (primarily letter triplet, or trigraph) frequencies. The software uses these tables to order the letters of the alphabet by most-likely frequency of occurrence, and displays the letters that the user is most likely to type next. If the desired letter is not in the initial set, users can touch an "Other" button to display the next most-likely letters. For small devices, users also need to touch buttons to display numbers and punctuation. The main advantage of a predictive keyboard is that its keys are much larger than the keys for a standard layout, given the same area in which to work. The key disadvantages are 1) the requirement to touch buttons to display characters, numbers, or punctuation not currently visible and 2) that the letters do not always appear in the same position, but instead appear in an order determined by their likelihood of use.

A third approach is handwriting recognition. A fast handwriter can produce about 25 wpm (Bailey), so handwriting is a feasible alternative to selecting letters from a standard layout, given perfect handwriting recognition. However, current handwriting recognition algorithms do not produce perfect recognition. Consequently, people who use a handwriting recognition system must constantly check the accuracy of recognition, which slows their writing considerably and often leads them to choose to stop using the handwriting recognizer and, instead, use the alternative method of an on-screen standard keyboard layout (Lewis, LaLomia, and Miranda, 1994).

The purpose of this study was to compare these approaches (standard keyboard, predictive keyboard, and simulated perfect handwriting recognition) considering user input rates and preferences. The results of this comparison can help PDA and PC designers decide which type of input method to emphasize in their interfaces.

Method

Participants

Six IBM employees participated in the experiment. Two were permanent employees, and the other four were graduate students interning at IBM. Three were male, and two were left-handed. Their reported typing speeds ranged from 25 to 60, and averaged 52.5. Three had used reduced-size keyboards previously, and those same three participants had previous experience using pen-based systems. Four participants had some PDA experience.

Materials

The participants used a Simon to input eight sentences with the Simon's built-in standard keyboard (typing), predictive keyboard (typing), or sketch pad (handwriting). The touch-sensitive screen on a Simon is 35 mm x 115 mm. The standard keyboard layout occupies a space 25 mm x 115 mm, with keys 9 mm by 5 mm. The predictive keyboard's area is 35 mm x 50 mm, with typing keys 13 mm x 13 mm. The drawing (writing) area of Simon's sketch pad is 33 mm x 83 mm.

I randomly selected 24 sentences from the Brown corpus (a collection of magazine and newspaper articles representative of contemporary English-language usage), and randomly distributed them among three sets of sentences. All sentences ranged in length from 90 to 110 characters. Three sets of addresses (each composed of eight addresses) contained addresses randomly selected from the 1994 Human Factors and Ergonomics Society directory. The final three stimulus sets included one sentence set and one address set, with the sentence-address set pairing determined by chance.

The other materials were an instruction sheet, a background questionnaire, a layout rating form for each input method, a keyboard layout attribute importance form, and a layout ranking form.

Procedure

I used a digram-balanced Greco-Latin design to counterbalance the presentation of both input method and stimulus set. This presentation method also balanced the pairing of input methods and stimulus sets (Lewis, 1993). Participants used all three input methods to enter both normal text (sentences) and addresses. They always entered the sentences first. The experimental design was within-subjects, with three independent variables (three types of input method, two types of text to enter, and eight trials per input method and text combination).

At the start of each experimental session, participants read the following instructions:

The purpose of this study is to evaluate your initial experiences with three ways of writing or typing on a small touchscreen. You can type and write with a stylus or with your finger, whichever you choose.

You will work with a different writing or typing method one at a time. Take as much time as you like to practice with each method. Then I will give you a sheet of paper with eight sentences or eight addresses on it. Please set everything up so you will be as comfortable as possible while you are typing or writing. When you are ready, start the first sentence. Try to be as fast and accurate as you can. I will time how long it takes you to type the sentence. We will do that for the other sentences or addresses, and then I will give you a rating form for the layout. (Don't make too much of the times that I write down for each sentence or address. Keep in mind that they are not equally easy to type.) After you complete the typing and writing for all three methods, I will give you two final evaluation forms to fill out. Then the experiment will be over.

Do you have any questions about the study? OK, let's start.

Participants completed a background questionnaire, then started the experiment. Each time they finished the sentences and addresses with an input method, they completed a form to rate the input method. After completing all input methods, they completed one form to rate the importance of various layout attributes, and a final form to rank (and discuss their reasons for ranking) the three input methods.

Results

Performance

The results of an analysis of variance on the entry rate (in characters per second, or cps) showed significant main effects for Input Method ($F(2,10)=44.8$, $p=.0001$), Text Type ($F(1,5)=4.9$, $p=.018$), and Trial ($F(7,35)=5.0$, $p=.001$). The Input Method by Text Type interaction was also significant ($F(2,10)=5.2$, $p=.029$), but none of the interactions with Trial were significant (Input Method by Trial: $F(14,70)=1.1$, $p=.369$; Text Type by Trial: $F(7,35)=1.6$, $p=.175$; Input Method by Text Type by Trial: $F(14,70)=0.9$, $p=.588$). The mean cps (and 95% confidence intervals) for each input method appears in Table 1 (and graphically in Figure 1). The mean cps for the Input Method by Text Type interaction appears in Table 2 (and graphically in Figure 2). In parentheses next to each tabled value for cps is the corresponding value in words per minute (wpm). Figure 3 shows the nonsignificant Input Method x Text Type x Trial interaction.

Because the 95% confidence intervals in Table 1 do not overlap, the results indicate that handwriting was significantly faster than the standard layout, which was significantly faster (in fact, over twice as fast) as the predictive layout. Figure 2 shows that, even though the interaction of Input Method and Text Type was significant, the interaction was due to differences in degree of difference between input speed for sentences and addresses, but there were no crossed effects. Entering addresses was always slower than entering sentences.

Table 1. Mean CPS and 95% Confidence Intervals for Input Method in CPS (WPM)

Input Method	Lower Limit	Mean	Upper Limit
Standard	1.10 (13.18)	1.18 (14.11)	1.25 (15.04)
Predictive	0.49 (5.84)	0.52 (6.23)	0.55 (6.62)
Handwriting	1.34 (16.07)	1.88 (22.62)	2.43 (29.16)

Note: The formula for converting cps to wpm is *wpm = cps * 60 sec/min * 1/5 words/characters.*

Table 2. The Input Method by Text Type Interaction in CPS (WPM)

	Type of Text	
Input Method	Sentences	Addresses
Standard	1.42 (17.02)	0.93 (11.21)
Predictive	0.60 (7.21)	0.44 (5.24)
Handwriting	1.97 (23.58)	1.80 (21.65)

Note: The formula for converting cps to wpm is *wpm = cps * 60 sec/min * 1/5 words/characters.*

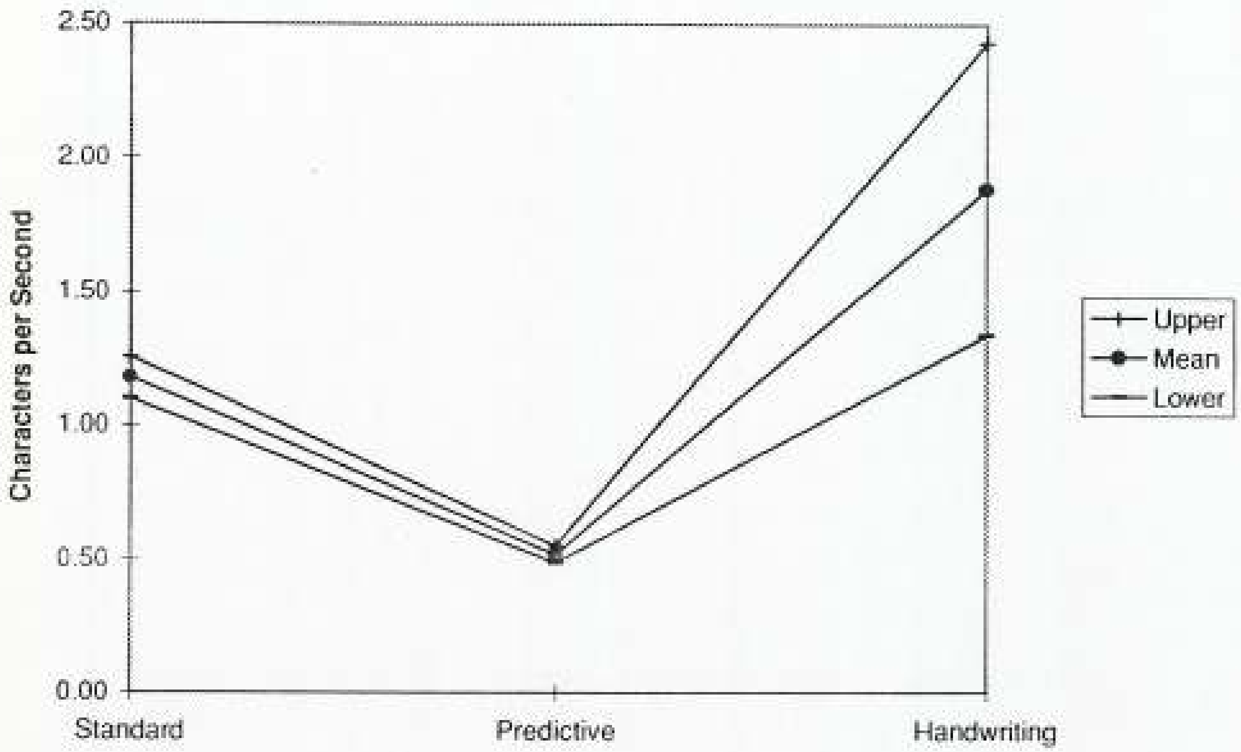

Figure 1. Mean CPS (with 95% confidence intervals) as a Function of Input Method

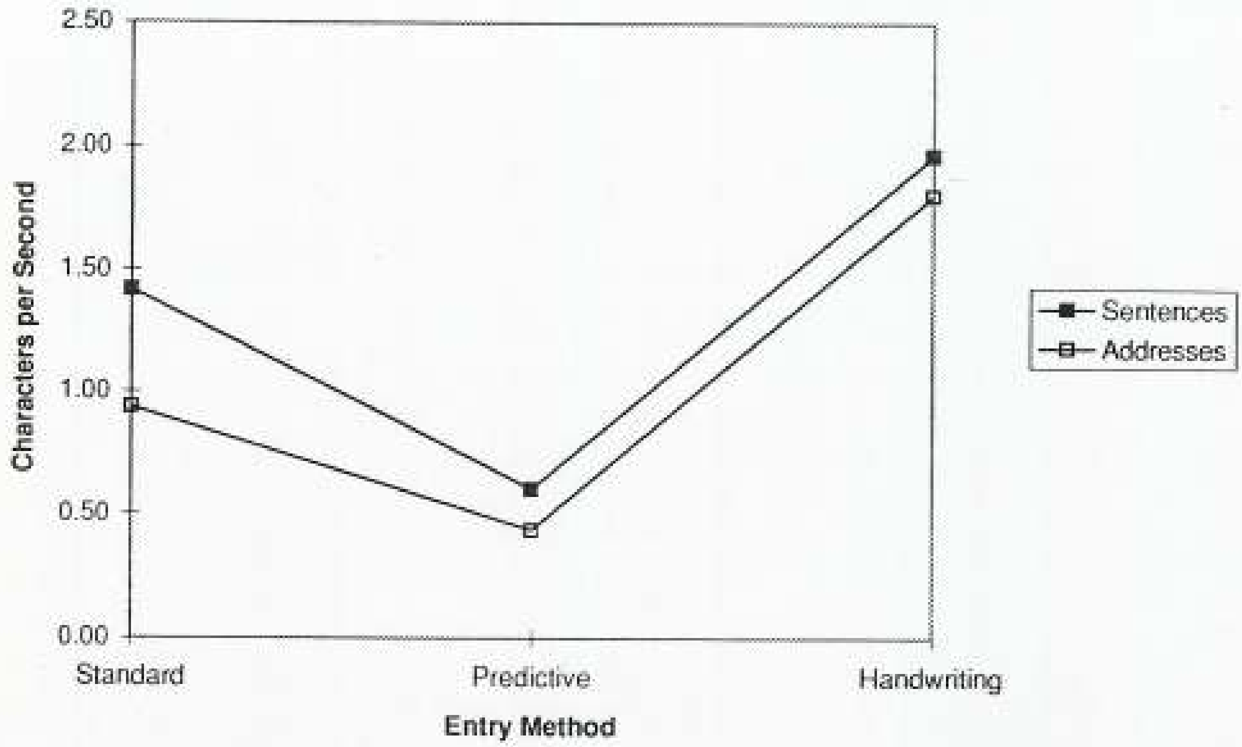

Figure 2. The Input Method x Text Type Interaction

Because learning usually follows an exponential function, any strong learning effect should show up as a steep, positively sloped line. Although the main effect of Trial was significant, the Input Method x Text Type x Trial interaction was not. As indicated by the nonsignificant interaction, Figure 3 does not support the hypothesis of a strong learning effect for any of the input methods.

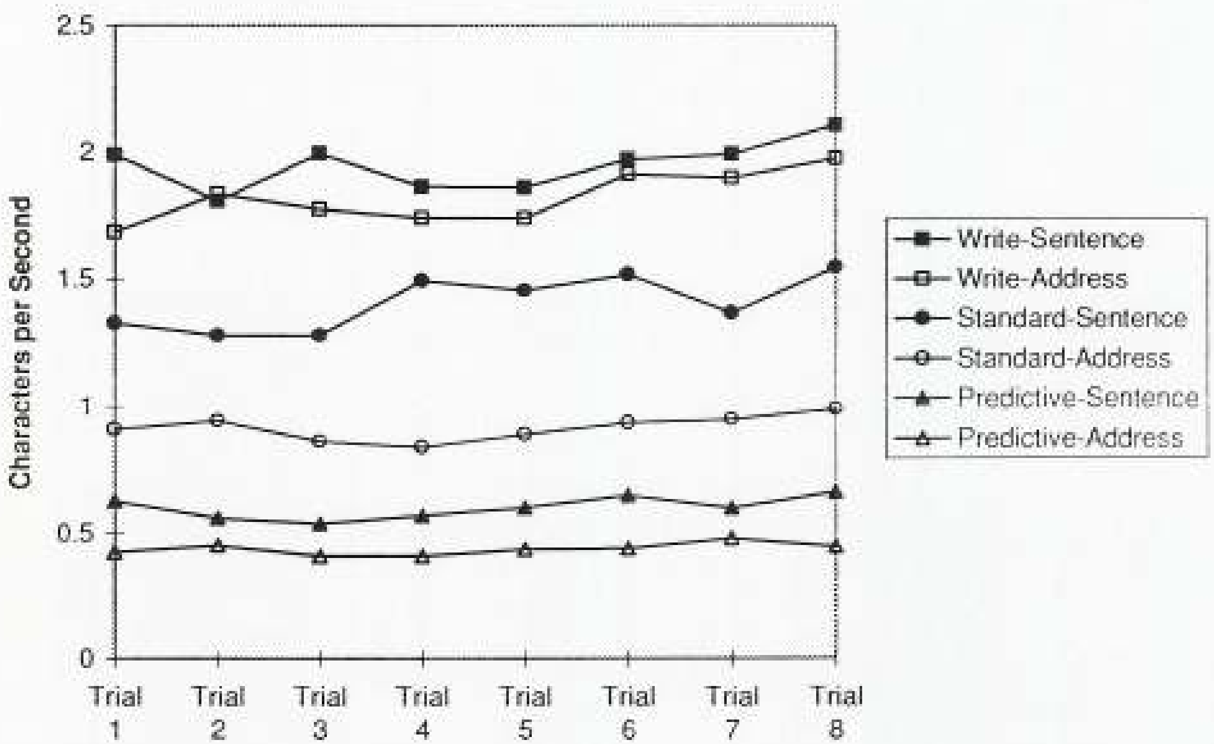

Figure 3. The Input Method x Text Type x Trial Interaction

Ratings

The rating questionnaires used 7-point bipolar scales, with lower ratings better than higher ratings. (See the appendix for copies of the rating questionnaires.) The average ratings for the input methods were significantly different ($F(2,10)=8.2$, $p=.008$). The average rating for the standard layout was 1.5, that for the predictive layout was 3.6, and the handwriting rating was 4.6. Multiple t-tests indicated that participants provided more favorable ratings for the standard layout than the predictive layout ($t(5)=4.6$, $p=.006$) and handwriting ($t(5)=4.3$, $p=.007$). The difference between the predictive layout and handwriting was not significant ($t(5)=1.0$, $p=.362$). Participants rated all the questionnaire attributes as important, with average importance ratings exceeding 6.0 for all attributes (except acceptability of key layout, which averaged 5.5).

Rankings

The mean ranks for the input methods were 1.33 for the standard layout, 2.17 for the predictive layout, and 2.50 for handwriting (a lower mean rank is better). A Friedman test indicated a marginally significant difference among the mean ranks ($X^2(2)=4.4$, $p=.10$).

Participant Comments

The main disadvantage of handwriting was the difficulty of handling a small device and writing comfortably. Also, the resolution and parallax of the display made it difficult to dot "i's" or cross "t's". The primary disadvantage cited for the predictive layout was the constantly changing character location. The main advantage cited for the standard layout was its familiarity. Table 3 presents a list of the reasons participants provided to explain why they ranked the input methods as they did.

Table 3. Participant's Reasons for Preferences

Participant	Choice	Method	Reason
1	1	Standard	Seemed most natural, fastest, most accurate.
	2	Predictive	More accurate than handwriting.
	3	Handwriting	Display too small, resolution too low. Parallax made it difficult to cross t's or write E's.
2	1	Predictive	It would be easier and faster than either of the others once I was used to it.
	2	Standard	It is what I have always used and is familiar.
	3	Handwriting	It was uncomfortable to print/write on such a small surface/area.
3	1	Standard	Easy to find letters -- always there and in same spot. Typing went much faster, and felt infinitely less complex.
	2	Handwriting	I like being able to write rather than type, but my messy handwriting came out even messier. I'm sure 20-30% of it or more is unreadable. Also, hard to orient the pad the best. Vertical comfortable to hold, but awkward space; horizontal space better, but harder to hold.
	3	Predictive	The keys were never in the same spot; spent much time trying to find the keys I wanted - either missed them first time around, had to hit "Other", "123", etc., or just trying to locate them on the screen. Neat idea, but not practical.

Table 3 (Cont.). Participant's Reasons for Preferences

Participant	Choice	Method	Reason
4	1	Handwriting	This was the easiest on my eyes, less visual search involved.
	2	Standard	This was particularly easy for me because I have much experience with that keyboard layout and knew where the keys were ahead of time, less thinking and mental workload involved.
	3	Predictive	It was disconcerting for me to have the location of the keys change so much because I tend to anticipate future keystrokes.
5	1	Standard	I am already familiar with the "QWERTY" arrangement and could find the targets quickly.
	2	Predictive	This is very novel -- but I was having to spend too much time searching for my "target" key each time I asked for "Other" -- may be a function of learning.
	3	Handwriting	I found the stylus to be fatiguing, and the result was not very readable. The pressure of my hand (fingertip) resting on the screen caused extra, unwanted lines.
6	1	Standard	I am familiar with the layout, therefore it is faster, easier, and no concentrating involved.
	2	Predictive	After 2 or 3 sentences I learned where the letters were, and I became more familiar with the procedure.
	3	Handwriting	The writing area was too small. The surface needs texturing so you don't slip.

Discussion

Participants indicated (with both rankings and ratings) that they preferred to use the standard keyboard. The average input rate for handwriting was fastest, but also much more variable than the other methods. Despite its speed, participants generally found it difficult to write comfortably and legibly on the small display. The input rate for the standard keyboard was more than twice the input rate for the predictive keyboard.

These results suggest that, for small devices, neither handwriting recognition nor predictive keyboards can effectively replace the standard keyboard layout. Until developers can invent something better, the standard keyboard layout should be the default input method for small devices. Even with perfect handwriting recognition, users will prefer to tap on a small standard keyboard unless the device's hardware design allows comfortable handwriting input with near perfect accuracy.

Experience in providing product demonstrations of Simon has shown that the predictive layout has a strong initial appeal, so it is reasonable to continue to include it in future devices. However, the predictive keyboard is markedly inferior to the standard layout for both sentence and address entry. It also appears that the problems participants experience using the predictive keyboard prevent it from being an input method that could benefit from practice. Uncertainty concerning letter location and the need to touch buttons to display numbers and punctuation cause user performance to be very slow with this input method. This is similar to the usability problems reported by Koester and Levine (1994) for word prediction interfaces. Although word prediction reduced the number of required keystrokes, "the cognitive cost of using word prediction had a major impact on the performance of these subjects" (Koester and Levine, 1994, p. 177). The current study showed a similar cognitive cost associated with using letter prediction.

Rather than investing additional development effort to obtain slight improvements in the predictive layout (see Lewis, Allard, and Hudson, 1994), it would make more sense to develop alternative layouts based on English-language digraph analysis (see Lewis, 1992b). A digraph-based layout takes advantage of the statistical characteristics of English letter distributions, but should be more learnable than a predictive layout because the letters are always in the same position.

The results of this experiment reinforce a finding from early competitive analyses of PDAs (Lewis, 1994). None of these input methods effectively replace the use of a full-sized, standard keyboard (such as a computer keyboard). Until the development of more effective small-device input methods, designers of PDAs should provide a computer version of the PDA's databases (with data upload from and download to the PDA) or an attachable full-sized keyboard to allow users to rapidly enter their information into the PDA.

References

Bailey, R. W. (1989). Human performance engineering: Using human factors/ ergonomics to achieve computer system usability. Englewood Cliffs, NJ: Prentice Hall.

Francas, M., Brown, S., and Goodman, D. (1983). Alphabetic entry procedure with small keypads: Key layout does matter. In Proceedings of the Human Factors Society 27th Annual Meeting (pp. 187-190). Santa Monica, CA: Human Factors Society.

Koester, H. H., and Levine, S. P. (1994). Modeling the speed of text entry with a word prediction interface. Transactions on Rehabilitation Engineering, 2, 177-187.

Lewis, J. R. (1992a). Literature review of touch-screen research from 1980 to 1992 (Tech. Report 54.694). Boca Raton, FL: International Business Machines Corp.

Lewis, J. R. (1992b). Typing-key layouts for single-finger or stylus input: Initial user preference and performance (Tech. Report 54.729). Boca Raton, FL: International Business Machines Corp.

Lewis, J. R. (1993). Pairs of Latin squares that produce digram-balanced Greco-Latin designs: A BASIC program. Behavior Research Methods, Instruments, & Computers, 25, 414-415.

Lewis, J. R. (1994). Competitive usability evaluation of Sharp Wizard OZ-9600, Hewlett-Packard HP-95LX, and Angler (Tech. Report 54.866). Boca Raton, FL: International Business Machines Corp.

Lewis, J. R., Allard, D. J., and Hudson, H. D. (1994). Predictive keyboard design study: Effects of different word populations, number of displayed letters, and number of trigraph tables (Tech. Report 54.846). Boca Raton, FL: International Business Machines Corp.

Lewis, J. R., LaLomia, M. J., and Miranda, S. (1994). Competitive evaluation of Newton using personal communicator scenarios (IBM Tech. Report 54.865). Boca Raton, FL: International Business Machines Corp.

Norman, D. A., and Fisher, D. (1982). Why alphabetic keyboards are not easy to use: Keyboard doesn't much matter. Human Factors, 24, 509-519.

Sears, A. (1991). Improving touchscreen keyboards: Design issues and a comparison with other devices. Interacting with Computers, 3, 253-269.

Trademarks

** Newton is a trademark of the Apple Corporation.

** Simon is a trademark of the BellSouth Corporation.

Appendix. Rating Questionnaires

Key Layout Rating Form (Standard)

Please rate the method you just used. Circle the number that best represents your judgment.

Easy to Find Letters	1--------2--------3--------4--------5--------6--------7	Hard to Find Letters
Easy to Type Fast	1--------2--------3--------4--------5--------6--------7	Hard to Type Fast
Easy to Type Accurately	1--------2--------3--------4--------5--------6--------7	Hard to Type Accurately
Easy to Learn Letter Locations	1--------2--------3--------4--------5--------6--------7	Hard to Learn Letter Locations
Easy to Type	1--------2--------3--------4--------5--------6--------7	Hard to Type
Key Layout Acceptable	1--------2--------3--------4--------5--------6--------7	Key Layout Unacceptable

Key Layout Rating Form (Predictive)

Please rate the method you just used. Circle the number that best represents your judgment.

Easy to Find Letters	1--------2--------3--------4--------5--------6--------7	Hard to Find Letters
Easy to Type Fast	1--------2--------3--------4--------5--------6--------7	Hard to Type Fast
Easy to Type Accurately	1--------2--------3--------4--------5--------6--------7	Hard to Type Accurately
Easy to Learn Letter Locations	1--------2--------3--------4--------5--------6--------7	Hard to Learn Letter Locations
Easy to Type	1--------2--------3--------4--------5--------6--------7	Hard to Type
Key Layout Acceptable	1--------2--------3--------4--------5--------6--------7	Key Layout Unacceptable

Rating Form (Handwriting)

Please rate the method you just used. Circle the number that best represents your judgment.

Easy to Write Fast	1--------2--------3--------4--------5--------6--------7	Hard to Write Fast
Easy to Write Accurately	1--------2--------3--------4--------5--------6--------7	Hard to Write Accurately
Easy to Write	1--------2--------3--------4--------5--------6--------7	Hard to Write
Writing Area Acceptable	1--------2--------3--------4--------5--------6--------7	Writing Area Unacceptable

Key Layout Attribute Importance Form

Please rate the importance of the following key layout usability features. Circle the number that best represents your judgment.

1. Ease of finding letters

Unimportant 1--------2--------3--------4--------5--------6--------7 Important

2. Ease of typing fast

Unimportant 1--------2--------3--------4--------5--------6--------7 Important

3. Ease of typing accurately

Unimportant 1--------2--------3--------4--------5--------6--------7 Important

4. Ease of learning letter locations

Unimportant 1--------2--------3--------4--------5--------6--------7 Important

5. Ease of typing

Unimportant 1--------2--------3--------4--------5--------6--------7 Important

6. Acceptability of key layout

Unimportant 1--------2--------3--------4--------5--------6--------7 Important

Are there any other key layout attributes that you think are important, and that you want to discuss with me?