2009-12-17

Design and Evaluation of Three Alternative Keyboard Layouts for a Five-Key Text Entry Technique

Barbara Millet
University of Miami, bmillet@gmail.com

UNIVERSITY OF MIAMI

DESIGN AND EVALUATION OF THREE ALTERNATIVE KEYBOARD LAYOUTS FOR A FIVE-KEY TEXT ENTRY TECHNIQUE

By

Barbara Millet

A DISSERTATION

Submitted to the Faculty
of the University of Miami
in partial fulfillment of the requirements for
the degree of Doctor of Philosophy

Coral Gables, Florida

December 2009

UNIVERSITY OF MIAMI

A dissertation submitted in partial fulfillment of
the requirements for the degree of
Doctor of Philosophy

DESIGN AND EVALUATION OF THREE ALTERNATIVE KEYBOARD LAYOUTS FOR A FIVE-KEY TEXT ENTRY TECHNIQUE

Barbara Millet

Approved:

____
Shihab S. Asfour, Ph.D.
Professor of Industrial Engineering

____
Terri A. Scandura, Ph.D.
Dean of the Graduate School

____
Murat Erkoc, Ph.D.
Assistant Professor of Industrial Engineering

____
James R. Lewis, Ph.D.
Senior Human Factors Engineer
IBM Corporation
Boca Raton, Florida

____
Eunji Lim, Ph.D.
Assistant Professor of Industrial Engineering

____
Arzu Onar, Ph.D.
Assistant Member, Biostatistics
St. Jude Children's Research Hospital
Memphis, Tennessee

MILLET, BARBARA (Ph.D., Industrial Engineering)
Design and Evaluation of Three Alternative Keyboard Layouts for a Five-Key Text Entry Technique. (December 2009)

Abstract of a dissertation at the University of Miami.

Dissertation supervised by Professor Shihab S. Asfour.
No. of pages in text. (293)

Despite the increase in popularity of handheld devices, text entry on such devices is becoming more difficult due to reduced form factors that limit display size, input modes, and interaction techniques. In an effort to circumvent these issues, research has found that five-key methods are effective for text entry on devices such as in-car navigation systems, television and gaming controllers, wrist watches, and other small devices. Five-key text entry methods use four directional keys to move a selector over an on-screen keyboard and an Enter key for selection. Although other researchers have described five-key character layouts using alphabetical order and predictive layouts based on digraph frequencies, there is considerable latitude in designing the rest of a comprehensive on-screen keyboard. Furthermore, it might be possible to capitalize on the relative strengths of the alphabetic and predictive layouts by combining them in a hybrid layout. Thus, this research examines the design of alternative keyboard layouts for five-key text entry techniques. Three keyboard layouts (Alphabetical, Predictive, and Hybrid) were selected to represent standard and less familiar arrangements. The analysis centered on a series of controlled experiments conducted on a research platform designed by the author.

In this work, when the immediate usability of three alternative keyboard layouts for supporting five-key text entry was investigated, results indicated no statistically significant differences in performance across the tested keyboards. Furthermore, experimental results show that following immediate usability, but still at the onset of learning, there was no overall difference in performance among the three keyboard layouts across four text types. However, the Alphabetical keyboard surpassed both the Predictive and Hybrid keyboards in text entry speed in typing Web addresses. The nonstandard keyboards performed superior to the Alphabetical keyboards in typing Words/Spaces and Sentences, but performed no better in typing Address strings than the Alphabetical.

Use of mixed effects modeling suggested that the longitudinal data was best fitted by a quadratic model. Text entry performance on all three layouts improved as a function of practice, demonstrating that participants could learn the unfamiliar layouts to complete text entry tasks. Overall, there was no indication that use of nonstandard layouts impedes performance. In fact, trend in time data suggests that the learning rates were greater for the nonstandard keyboards over the standard layout. Overall, participants preferred the Hybrid layout. In summary, this dissertation focused on creating and validating novel and effective five-key text entry techniques for constrained devices.

ACKNOWLEDGMENTS

I would like to thank my advisor Dr. Shihab S. Asfour for his faithful support and guidance. Thank you, Dr. Asfour, for always putting your trust in me. I would not be where I am today without the time and effort you provided throughout my graduate studies.

I also would like to extend my appreciation to Dr. James R. Lewis for his mentorship. Dr. Lewis introduced me to this dissertation topic and to text entry research. I cannot imagine being able to complete this work without his guidance. Thank you, Jim, for your insights, comments, suggestions, edits, and for your friendship.

Special thanks to Dr. Arzu Onar for her time and effort in improving this work. Thank you, Dr. Onar, for your significant contribution to the statistical aspects of this work and for your thoughtful reviews.

Thanks also go to my other committee members, Dr. Murat Erkoc and Dr. Eunji Lim for reviewing my work and for their helpful suggestions.

I am grateful to my husband, Philippe, for his love, support, patience, and help. I am looking forward to our lives together post dissertation.

I am indebted to my family, especially to my mother and to my aunt, for believing in me and for motivating me to complete this work.

Lastly, special thanks to Motorola's Design Integration group (Plantation, FL) for granting permission to run my experiments in their lab and to my participants for volunteering their time.

To my family

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Chapter 1. Introduction

There is a strong demand for mobile computing devices such as PDAs, in-car navigation systems, and cellular phones. Indeed, the mobile domain is overwhelmed by consumer demands for micro devices that offer full functionality, including some tasks that require text entry. By definition, micro devices have limited physical space for text input and output, and there is still a need for efficient text entry methods that can work within these constraints.

To support text entry on smaller devices, some researchers have investigated text entry using only five keys. Five-key text entry techniques rely on four physical directional keys to move a selector over an on-screen keyboard and a fifth key for selection. Although clearly less efficient than a full-sized standard keyboard, five-key methods require relatively little space and are easy to learn (Wobbrock, Myers, & Rothrock, 2006).

Several researchers have conducted studies of five-key text entry techniques using various keyboard layouts, such as alphabetic, QWERTY, and predictive. Bellman and MacKenzie (1998) proposed a hybrid layout, which combines fixed and predictive layouts, but did not provide any analyses of its efficacy. Such a combination could, however, prove advantageous for extremely small user interfaces. Thus, the focus of this work is a "hybrid" layout (one that combines a full fixed and limited predictive character set), with comparison to a fixed and a predictive keyboard designed with similar constraints. The fixed keyboard is alphabetical, and the predictive keyboard uses Bellman and MacKenzie's Fluctuating Optimal Character Layout (FOCL) strategy. The FOCL strategy relies on letter-pair (digraph) probabilities to dynamically rearrange letters (based on the most recently selected letter) for the purpose of minimizing selector movement.

1.1 Motivation

Despite the increase in popularity of handheld devices, text entry on such devices is becoming more difficult due to reduced form factors that limit display size, input modes, and interaction techniques. In an effort to circumvent these issues, research has found that five-key methods are effective for text entry on devices such as two-way pagers, wrist watches, and other small devices. However, the requirements for multiple keystrokes per character and high visual scan time prevent these methods from being practical in many contexts. Innovative solutions to these problems, focusing on predictive five-key text entry techniques, are in the focus of this research, with the primary goal to create and validate efficient five-key text entry methods.

1.2 Contributions

The theoretical and practical contributions of this dissertation include: (a) Design of alternative keyboard layouts that promote novice and expert usability. (b) Empirical validation of the proposed layouts for both immediate usability and usability after multiple trials of experience. (c) Development of a predictive model for expert text entry performance using mixed effects modeling techniques. (d) Development and validation of an extensible platform for the development of prototypes of alternative selection-based text entry schemes and their empirical evaluation.

Chapter 2. Background

This chapter discusses the evaluation methods for text input techniques and briefly reviews related literature.

2.1 Empirical Evaluation

Evaluating the effectiveness of a new text entry technique is an important aspect of the development of these methods. Most empirical evaluations of input devices are comparative (MacKenzie & Soukoreff, 2002b). Comparative evaluations allow the assessment and comparison of new text entry techniques against current or existing methods. Comparative evaluations can also be used to compare alternative design configurations of a specific input method.

Detailed planning is required when conducting such an evaluation, as confounding factors could adversely affect its repeatability and validity (MacKenzie & Soukoreff, 2002b). Empirical evaluations are conducted in constrained, artificial environments to control confounding factors. These constrained environments (i.e. laboratories) facilitate the isolation and accurate measurement of the behaviors of interest. Additionally, to ensure validity of the results, empirical evaluations must be representative of actual user behavior (MacKenzie & Soukoreff, 2002b). The following sections cover important issues in conducting valid and useful evaluations of text entry techniques.

2.1.1 Evaluation Task

In a typical text entry evaluation, the experimental task is the input of text using the entry method(s) under study. Generally, there are two types of tasks: text creation and text copy. In a text creation task, the participant memorizes or composes the source text so there is no need to refer to an external copy (MacKenzie & Soukoreff, 2002b). This differs greatly from a text copy task, where the participant is given the text to enter (as in Butts & Cockburn, 2001; Curran, Woods, & Riordan, 2006; Dunlop & Crossan, 2000; James & Reischel, 2001; MacKenzie & Zhang, 1999; and MacKenzie & Zhang, 2001).

Although text creation more closely mimics real world usage, text copy tasks are commonly used in empirical evaluations. This is mainly because behavioral aspects in composing text, such as "pondering", are difficult to measure (MacKenzie & Soukoreff, 2002b). Additionally, in text creation tasks it is difficult to ascertain whether an error was committed or not, as it is impossible to know what the participant intended to enter. Consequently, text copy (transcription) is the favored approach in laboratory settings when evaluating text entry methods (Wobbrock, 2007).

2.1.2 Error Handling

Another important consideration for a text entry evaluation is the determination of error handling. There are very different ways for handling errors in a text entry study. Some experimenters disallow input of incorrect characters (as in Evreinova, Evreinov, & Raisamo, 2004), while others disable the backspace (as in MacKenzie & Zhang, 1999; Matias, MacKenzie, & Buxton, 1996). An alternative method is to allow participants to correct errors, an approach referred to as unconstrained text entry (Wobbrock, 2007).

Typically during a unconstrained text entry experiment, short strings (input) are presented, one at a time, to participants who enter the strings (transcribed) using the text entry method under study (MacKenzie & Soukoreff, 2002b). Comparison of this pair of strings (input vs. transcribed) reveals any errors that occurred when using the text entry method. Participants receive instruction to enter the string as "quickly and as accurately as possible" and can correct mistakes (as in Curran et al., 2006). This process is usually repeated over several trials.

In addition to the input and transcribed strings; the data also includes the events from the input stream (the sequence of actions taken by the participant to generate the transcribed string), usually captured using log files. The data is then parsed and analyzed to obtain the appropriate measures to report for the specific text entry technique under investigation (Wobbrock, 2007).

2.1.3 Prototypes and Data Collection

When designing an empirical evaluation the test equipment to use must be considered. Ideally, the devices used in the evaluations should be high fidelity prototypes. In practice, a working prototype might not exist and development resources needed to create it may not be available (Silfverberg, MacKenzie, & Korhonen, 2000). An alternative approach is to use software-based simulators (Bellman & MacKenzie, 1998; Butts & Cockburn, 2001; Dunlop & Crossan, 2000, MacKenzie, 2003; MacKenzie, Kober, Smith, Jones & Skepner, 2001; MacKenzie & Zhang, 1999).

These simulations provide experimenters control over the experimental conditions and allow automation of the test protocol. However, this approach reduces the external validity of study results, because participant interaction with the simulation is not necessarily representative of how users will interact with the actual devices (Sirisena, 2002). On the other hand, a principal advantage of this approach is the ability to automate data collection, therefore minimizing time-consuming and error-prone hand tabulation (Wobbrock, 2006).

2.1.4 Initial Performance and Longitudinal Leaning

Maximizing the ultimate performance of an average user is a primary design goal when developing new text entry methods (Zhai, Kristensson, & Smith, 2005). However, it is critical to also consider the immediate usability of any new input technique because new users are not likely to accept any strategy that requires considerable learning (MacKenzie & Zhang, 1997). Empirical investigations support the exploration of text entry performance at these various levels of user skill.

Immediate usability captures a user's initial experience with an interface. In doing so, performance can be evaluated at the onset of learning. Measuring immediate usability, however, requires careful control of the participants' exposure to the interface (as in Cockburn & Siresena, 2003; Isokoski, 1999; Koltringer & Grechenig, 2004; MacKenzie & Zhang, 2001).

Expert evaluations require longitudinal studies (as in Bellman & MacKenzie, 1998; Isokoski & Raisamo, 2004; Lyons, Plaisted, & Starner, 2004; Wigdor & Balakrishnan, 2004; Wobbrock et al., 2006; Zhai, Sue, & Accot, 2002). In longitudinal studies fewer people participate, but they use the device(s) over extended periods of time. Longitudinal studies are costly and time-consuming because they require numerous experimental sessions over a relatively long period of time. However, these studies are necessary to establish a learning curve for a particular text entry method and to estimate its optimal performance (Sirisena, 2002).

2.2 Measures of Text Entry

2.2.1 Performance

There are numerous empirical measures of text entry performance. However, the two primary metrics are speed and accuracy (MacKenzie & Soukoreff, 2002b). These two measures have extensive applicability to a variety of text entry methods.

2.2.1.1 Speed

Speed is a key feature for most text entry techniques (Wobbrock, 2007), typically measured as characters per second (CPS) and calculated as follows:

$$CPS = \frac{|T|}{S}$$

where *T* is the length of the final transcribed string, and *S* is the time, in seconds, taken to enter it. Here, *T* contains all characters in the string, such as letters, spaces, numbers, punctuations and other printable characters (Wobbrock, 2007). *T* does not include backspaces. Typically, the calculation for CPS is modified to the following:

$$CPS = \frac{|T|-1}{S}$$

Here, the minus one in the numerator accounts for when timing for text entry begins. Because *S* is measured from the entry of the first character to the entry of the last character, the preparation time leading to the input of the first character is not captured. As a result, the character count is decremented by one to allow for a more accurate measure of entry speed. Further, the standard for text entry speed is to report words per minute (WPM), derived from CPS, as follows:

$$WPM = CPS \times 60 \times \frac{1}{5}$$

Therefore, to obtain WPM, multiply CPS by 60 seconds per minute and then divided by 5, which is the average number of characters in a word, including spaces (MacKenzie & Soukoreff, 2002b).

Alternative values have been proposed for this parameter, based on the analysis of representative text in a language corpus. For example, Dunlop and Crossan (2000) calculated a value of 5.98 for average word length including final space, based on their examination of a comparative text document. Additionally, James and Reischel (2001) also used the 5.98 value when conducting their study that compared model predictions to actual predictive (T9) and multi- tap performance.

2.2.1.2 Accuracy

In contrast to calculating entry speed, measuring accuracy is more challenging. A simple measure of error rate is to obtain the number of characters in error as a percentage of the length of the presented string (MacKenzie & Soukoreff, 2002b). However, a more thorough analysis involves understanding what type of errors occurred.

MacKenzie & Soukoreff (2002b) explain that there are four basic error types: 1. Substitution (entering an incorrect character), 2. Insertion (entering an extra character), 3. Omission (omitting a character), and 4.Transposition (reversing adjacent characters). However, when conducting unconstrained text entry experiments two new categories of errors are created. First, there are errors that are committed and then corrected. Second, there are errors that are committed but go unnoticed by the participant, and therefore remain in the transcribed string. In 2003, Soukoreff and MacKenzie introduced error metrics for calculating corrected and uncorrected error rates. In their framework, participant keystrokes are delineated into four groups:

- Correct (C) keystrokes= any keystroke not in error
- Incorrect and Not Fixed (INF) keystrokes= undetected errors that remain in the transcribed string
- Incorrect but Fixed (IF) keystrokes= errors that are committed and later corrected
- Fixed (F) keystrokes= any keystroke that performs a correction (e.g. backspace)

Note that this categorization requires the analysis of the input string, the transcribed string, and the input stream (Soukoreff & MacKenzie, 2003b). The following sections describe several resulting performance measures.

2.2.1.2.1 Uncorrected Errors

Uncorrected errors are those that remain in the in the transcription. The Uncorrected Error Rate is:

$$UncorrectedErrorRate = \frac{INF}{C + INF + IF} \times 100\%$$

2.2.1.2.2 Corrected Errors

Corrected errors are any characters that were backspaced during entry (Soukoreff & MacKenzie, 2003b). Corrected errors do not appear in the transcribed string and requires the analysis of the input stream. The Corrected Error Rate is:

$$CorrectedErrorRate = \frac{IF}{C + INF + IF} \times 100\%$$

2.2.1.2.3 Total Error Rate

The Total Error Rate combines the Corrected and Uncorrected Error Rates:

$$TotalErrorRate = \frac{IF + INF}{C + INF + IF} \times 100\%$$

2.2.2 Efficiency

Keystrokes per character (KSPC) is a well-known model-based method for quantifying the efficiency of a text entry technique, using the average number of keystrokes needed to produce each character with a specific text entry method (MacKenzie, 2002a; Wobbrock, 2007). Calculating KSPC requires the use of a language model and keystroke data. An estimate of the average KSPC is:

$$\overline{KSPC} = \frac{\sum (K_c \times F_c)}{\sum F_c}$$

where K_c is the number of keystrokes needed to enter character c and F_c is the frequency count for c in the corpus (MacKenzie, 2002a). Note the equation presented above is useful only if the entry of each character depends only on that character. If the keystrokes depend on the previously entered character, a more intricate formula is necessary.

The use of KSPC as a measure of efficiency assumes expert behavior (use of the shortest path without errors). Limitations of KSPC are that it measures only a single aspect of performance and does not consider the layout of numbers, punctuations and other special characters. Additionally, the KSPC measure is device dependent and therefore cannot be used to compare the efficiency between different text entry methods (Soukoreff & MacKenzie, 2003b)[1].

---

[1] However, the comparisons of KSPC are legitimate as long as there are no changes in the hardware used for text entry.

2.2.3 Subjective Measures

A participant's impression can be indicative of how well a text entry method performs and future usage of the method. Thus, it is important to collect participants' impressions of the text entry method under study, as subjective responses to an interface allows for interpretation of the quantitative measures gathered throughout the experiment. Typically, qualitative measures are collected in the form of questionnaires administered at the end of the evaluation. Standardized qualitative tests, such as the NASA Task Load Index, which is a subjective workload measurement tool (Hart & Staveland, 1988), are available in the literature. Similarly, there are numerous textbooks that guide questionnaire design (MacKenzie & Soukoreff, 2002b).

2.2.4 Learnability

Because skilled text entry involves learning, it is important to understand the effects of learning on user performance with any new text entry method. Learning, however, has different effects on entry rate and error rates (Isokoski, 2004). Error rates are typically very high initially, but then either stay the same or quickly fall to more tolerable levels. Entry rates, however, tend to improve following the power law of learning (Isokoski, 2004).

The power law of learning gives entry time as a function of the amount of practice. The learning curve can be approximated (De Jong, 1957; Card, English, & Burr, 1978) as follows:

$$T_N = T_1 \times N^{-\alpha}$$

where T_1 is the entry time on the first trial, T_N is the predicted entry time on the N^{th} trial, N is the session number, and α is an empirically determined constant. Accordingly, the "ease of learning" of a text entry method can be expressed by T_1 and α (Card, English, & Burr, 1978). These numbers are determined empirically by regressing log(T_N) on log(N). Although the curve tends to fit measured data well (as in MacKenzie & Zhang, 1999; McQueen et al., 1995), the power law of learning assumes that performance will improve indefinitely. This is not the case for text entry methods, however, where entry rates eventually reach an upper limit.

Isokoski and MacKenzie (2003) explain that the virtue of any text entry method depends on the attainable entry rate. As a result expert performance and the effort required to attain it, are generally the most critical characteristics of a text entry method (Isokoski & MacKenzie, 2003). Therefore, investigating performance improvements of the Hybrid five-key text entry method, introduced herein, is central theme of this study.

2.3 Modeling

Predictive models are used to estimate metrics of human performance. Such models allow the exploration of text entry designs without device implementation or measures derived from empirical evaluations (MacKenzie, 2003). In text entry research, there are two main approaches to predictive modeling: Keystroke Level Modeling (KLM) and Fitts' Law (a movement time model). These models are mainly used to predict expert, error-free performance and require the use of a language model to predict text entry rate.

A language model predicts the probabilities of characters or words in a language of interest. Language based models are derived from large samples of text (or corpus). Therefore, the quality of any language model is affected by the text on which it was based (Soukoreff & MacKenzie, 2003a). In fact, all language models are limited in that input modalities are not considered and the editing process used is not reflected (Soukoreff & MacKenzie, 2003a).

2.3.1 Keystroke Level Modeling

KLM is a simple model for the time it takes an expert user to complete a task, without error, with a given method (Card, Moran, & Newell, 1983). The model is based on the following four physical-motor operators:

- Keystrokes (K),
- Pointing (P),
- Homing (H),
- Drawing (D),

and one mental operator (M), plus a system response operator (R). So, execution time is the sum of each of the operators (Card, Moran, & Newell, 1980).

$$T_{Execute} = t_K + t_P + t_H + t_D + t_M + t_R$$

Given the task and the method, the KLM uses duration estimates, either in the form of a single value or a parameterized estimate, of these operators to predict task completion time for an expert user (John, 2003). In practice, some of the operators are omitted or repeated depending on the task and method under study (MacKenzie, 2003).

Since its introduction, KLM has been used for predicting text entry performance. For example, Detweiler, Schumacher, and Gattuso (1990) proposed KLM models for entering alphabetic strings on a telephone keypad, using four two-key methods (Top Row, Same Row, Modal Position, and Modified Modal Position) and a key repeat method. In using Two-key methods, the user presses two keys to specify a letter. With the multi-tap approach, the user presses a key multiple times to get a specific letter. The models proposed generally included four components for the task. The predicted text entry times were of 11.96, 13.72, 14.58, 13.78 and 13.50 seconds for Key Repeat, Modified Modal, Modal, Same Row, and Top Row, respectively. In a user study of the five methods, text entry times were of 12.38, 12.50, 14.81, 14.18 and 13.50 seconds for Key Repeat, Modified Modal, Modal, Same Row, and Top Row, respectively. They concluded that their models matched the empirical data well.

Koester and Levine (1994) have also used KLM to predict text entry performance for physically challenged users in a word prediction system. In their model, the unit task was entry of a single word, which was accomplished through a series of letter and word list selections. They explain that each of these selections involved cognitive, perceptual and motor component actions. Their model was made of two parameters, based on a linear combination of key presses and list search actions, and was shown to predict word entry times with an average error of 16%. The model they developed appeared to be more accurate for able-bodied participants compared to those with spinal cord injuries.

Dunlop and Crossan (2000) proposed keystroke level models to evaluate a predictive text entry method (similar to T9), a word completion text entry method (similar to the predictive text entry but extended with automatic word completion), and multi-tap.[2] The models predicted the time taken by an expert user to enter an alphabetic phrase without error. To develop their models, they used three timing components, based on KLM. These components were defined as keying time for button presses, homing time for hand movement, and mental preparation for action execution. Dunlop and Crossan's models predicted text entry rates of 14.9, 17.6, and 7.7 WPM for the multi-tap,

[2] See section 2.4.2 for a description of these methods.

the predictive method, and the word-completion method, respectively. Typing speeds for multi-tap and T9 have been reported to be around 10 and 20 WPM, respectively (Silfverberg, 2007).

2.3.2 Fitts' Law

Fitts' Law is a predictive model of human movement. Specifically, it predicts the time for hand movements to a target. Fitts' Law (1954) states that the time to acquire a target is a function of the distance to the target and the size of the target. Fitts' equation for movement time (MT) prediction is:

$$MT = a + b \times \log_2\left(\frac{A}{W}\right)$$

where a and b are regression coefficients determined through empirical test, A is the amplitude or distance from the starting point to the target, and W is the width of the target. This model is particularly powerful in predicting rapid aimed movements (MacKenzie, 2003).

In text entry, Fitts' Law is used to predict the time to enter each character from the language of interest. To build a text entry prediction model using Fitts' Law requires information regarding the position and size of keys, the letter assignment to the keys, and digram probabilities of the target language (MacKenzie, 2003). A disadvantage in using Fitts' Law to develop models is that behavior is explained by discrete actions. However, models for analyzing human-machine interactions need to deal with sequential, integrated, behavior rather than discrete actions (John, 2003). Even so, the validity of Fitts' Law for general tapping tasks is undisputed in the literature (Zhai et al., 2005).

Therefore, Fitts' Law remains the preeminent model for pointing device (stylus) research (MacKenzie, 2003).

Lewis (1992) was the first to use Fitts' Law and digram probabilities to model stylus keyboard performance (Zhai et al., 2005). Soukoreff and MacKenzie (1995), followed Lewis' effort and developed a theoretical model to predict upper and lower bounds on text entry speed using a stylus and a soft keyboard. Their model used the Hick-Hyman Law for choice reaction time, Fitts' Law for rapid aimed movements, and linguistic tables of English digrams. The model predicted a QWERTY text entry rate of 8.9 WPM for novices and 30.1 WPM for experts.

Based on the work of Soukoreff and MacKenzie (1995), Silfverberg, MacKenzie, and Korhonen (2000) developed prediction models, using a movement model based on Fitts' Law and letter digraph probabilities, for multi-tap and T9. Their models predict a one-handed thumb text entry rate of 20.8 WPM for Multi-Tap (without a time-out button) and 40.6 WPM for T9.

2.3.3 Fitts' Law and KLMs

Given the advantages and disadvantages of the two modeling approaches presented in the preceding sections, text entry researchers have pooled both modeling approaches to develop a combined model that considers motor and non-motor components in text entry (as in Dunlop & Masters, 2007; Pavlovych & Stuerzlinger, 2004). For example, Dunlop and Masters proposed a combined model of text entry performance based on keystroke modeling with key timing information derived from Fitts' Law. They investigated two predictive text entry methods (similar to T9). One of the methods was a five-key predictive text entry method with four alphabetic keys and a combined space/next key.

The other method used nine-keys for text entry. Their model predicted expert text entry rates from 20.2 to 23.0 WPM for the nine-key predictive keypad and 17.04 to 22.3 for the five-key predictive keypad. In a user study, participant entry rates were 21 and 12 WPM for the nine-key and five-key predictive keypads, respectively. Although the five-key text entry rates were slower than predicted, Dunlop and Masters explained that the model predicted expert five-key text entry rates and the participants of the study were novices to the five-key input method. Further, the empirical results for the five-key method were in line with novice nine-key user text entry rates of around 11 WPM (as reported in James & Reischel, 2001).

2.4 Mobile Text Entry Techniques

The prevailing text entry method for desktop computing is still the QWERTY keyboard (Silfverberg, 2007). The QWERTY layout was designed by Sholes, Glidden, and Soule in 1868. The QWERTY arrangement places many commonly adjacent letter pairs on the opposite sides of the keyboard. The original design goal was to minimize mechanical jamming, but had an unintended consequence of facilitating the frequent alternation of the left and right hand, which contributes to rapid touch typing (Zhai, Hunter, & Smith, 2002). Also, the full-size QWERTY keyboard is unambiguous in that each letter has a dedicated key. Therefore, each keystroke generates one character of text (KSPC= 1.0). Touch-typing speeds on desktop QWERTY keyboards range from 20 WPM to over 60 WPM for expert typists (MacKenzie & Soukoreff, 2002b). As a result, the standard QWERTY keyboard is considered to be the gold standard against which to assess the efficiency other input methods (Lewis, Commarford, Kennedy, & Sadowski, 2008).

Physical size is a major factor for mobile devices. To be competitive, consumer products must be as small as possible. A full-size QWERTY is familiar to many people, but takes up significant space even in reduced form. The necessary reduction in the size or number of keys available on mobile devices (due to space and weight considerations), means touch-typing is usually not practical or possible.

In mobile computing a full-size QWERTY keypad is not practical, which has driven the development of alternate text entry techniques. Given that mobile keypads often have fewer keys than characters in a language, more than one keystroke (KSPC$\geq$ 1.0) is usually required for each character (MacKenzie, 2002b).

The main point of differentiation between the various mobile text entry techniques is the type of input used. Most methods are either key-based or stylus/touch based (MacKenzie & Soukoreff, 2002b; Sirisena, 2002). In this review, only a few key-based methods (telephone keypads, small physical keyboards, and text selection) and one stylus/ touch based method (soft keyboards) are presented. Although other types of input methods exist, such as chording keyboards, speech recognition, handwriting recognition, and gesture-based input, these methods are not applicable given the constraints of the devices considered in this research. For comprehensive reviews of mobile text entry methods, see MacKenzie and Soukoreff (2002b) or Lewis et al. (2008).

2.4.1 Mini Physical Keyboards

Many recent handheld computers and smart phones are equipped with a small physical keyboard. The most used layout is the familiar QWERTY (see Figure 2-1), which allows for skill transfer from desktop computing. Mostly users hold the device with two hands and type with their thumbs (Silfverberg, 2007). As can be noted, reduction in the size keys available on mobile devices means touch-typing is usually not possible. Regrettably, there have been relatively few studies of small physical keyboards in the literature.

(a) (b) (c)

Figure 2-1 Examples of QWERTY Keyboards for the (a) Motorola Q9, (b) Nokia E62, and (c) Blackberry Curve

MacKenzie and Soukoreff (2002a) introduced a Fitts' Law based model of two-thumb text entry on small QWERTY keyboards. Their model predicted expert text entry as high as 60.74 WPM. Clarkson, Clawson, Lyons and Starner (2005) conducted a longitudinal study of mini-QWERTY keyboards. Initial typing rates were 31 WPM and after twentieth typing session the average was 60 WPM.

Matias et al. (1996) proposed a QWERTY-like keyboard that was reduced in size but leveraged touch typing skills. The Half-QWERTY keyboard, developed by Matias Corporation (www.halfqwerty.com), is a standard QWERTY keyboard split in half (see Figure 2-2). The Half-QWERTY displays all the keys used by one hand at one time. When the space key is pressed, the missing characters are mapped onto the keys. The Half-QWERTY keyboard has the advantage of eyes-free and one-handed operation (Sirisena, 2002), but is still relatively large for use in handheld computing.

Figure 2-2 Half-QWERTY Keyboard

2.4.2 Telephone Keypads

The standard input method for mobile phones is a 12-key keypad (see Figure 2-3). The keypad consist of the numbers 0-9 and two additional keys (# and *). The letters a-z are distributed across keys 2-9 in alphabetical order. The character layout is similar on most mobile phones, as it based on an ISO standard (Butts & Cockburn, 2001; Silfverberg et al., 2000; MacKenzie & Soukoreff, 2002b). However, the placement of the space key (on 0 in Figure 2-3) varies among phones (MacKenzie & Soukoreff, 2002b).

Figure 2-3 Standard 12-Key Mobile Phone Keypad

The 12-key keypad poses a challenge for text input since most languages have at least 26 letters. Given that there are more letters than keys, each key has three to four letters. As a consequence, text input using the standard keypad is ambiguous because there is uncertainty as to the intended character when a user presses a key.

The disambiguation techniques employed is major point of differentiation among the text entry methods that use the telephone keypad. The two main approaches include multi-tap (pressing a key multiple times to get a specific letter) and predictive (pressing keys just once, and then using predictive disambiguation to indicate the most likely word).

2.4.2.1 Multi-tap

With the multi-tap method, the user presses each key one or more times to specify the intended character. For example, the user would have to press the 4 key once for the “g” character, twice for the “h” character, and three times for the “i” character, thereby increasing the number of keystrokes required to enter a character. MacKenzie calculated that entering English text using the multi-tap method requires, on average, 2.034 keystrokes per character (2002a). Text entry rates are typically around 10 WPM (Silfverberg, 2007). Noticeably, the multi-tap method is tedious and inefficient (MacKenzie, 2003).

Additionally, the multi-tap method suffers with problems of segmentation. Segmentation problems occur when a character is on the same key as the previous character, such as when ‘e’ and ‘f’ are entered consecutively. Successive key presses are segmented using either a timeout or a dedicated kill key. Despite the problems identified for the multi-tap method, this method is still the most common text input method for mobile phones (MacKenzie & Soukoreff, 2002b; Lewis et al., 2008).

An alternative approach to the multi-tap method is to allow all or some characters to be entered with a single keystroke and to perform disambiguation by some other means. Collectively, these methods are predictive text input methods, commonly referred to as one-key with disambiguation.

2.4.2.2 Predictive Text Entry

Predictive text entry with a telephone keypad increases the efficiency of text entry to around one keystroke per character (MacKenzie, 2002a), if the prediction works well. To provide the one key press per letter feature, keystroke sequences are compared against a dictionary, containing word probabilities, to determine the desired word. For any given keystroke sequence, the system retrieves a list of words that could be entered with that specific sequence. For example, the key sequence 8-4-3 (see Figure 2-3) can represent "the", "tie", and "vie". The list provided to the user appears in descending order of word probabilities, with the most probable word presented first. The user can scroll through the list of predicted words by pressing the next button.

Predictions are made after each keystroke, not including the space key. Therefore, for each word of *n* length, *n* predictions are made (Sirisena, 2002). Given that each individual keystroke is ambiguous, predictions made for a key press can change based on subsequent key presses. This could be confusing for novice users as they cannot connect what appears on the display with what they want to enter (Sirisena, 2002). However, James & Reischel (2001) point out that these changes are often unnoticed because the user's attention tends to be on the keypad rather than the display during entry. Unfortunately, the changing predictions compounded by the fact that users are focusing on the keypad and not the display makes error correction difficult as it is not obvious at what point the error occurred.

Another limitation of this disambiguation technique is that some words may not be in the dictionary, in which case disambiguation fails. When this happens, the user typically has the option to use multi-tap to enter the word (and might be able to add the new word to the dictionary for subsequent use). A well-known example of predictive text input on mobile phones is T9 (Silfverberg, 2007), developed by Tegic Communications Inc[3] and reviewed extensively in the literature.

Silfverberg, MacKenzie, and Korhonen (2000) proposed mathematical models, based on Fitts' Law and letter digraph probabilities, for multi-tap and predictive text entry. The models predicted expert entry rates of 27 and 45.7 WPM for multi-tap and for predictive text entry, respectively.

James and Reischel (2001) undertook a study to compare model predictions to actual multi-tap and predictive text entry rates. In their study, they also explored how text entry rate differed between novice and expert users. Study results for novices showed that predictive text entry (9 WPM) was slightly faster than multi-tap (8 WPM). For experts, predictive text entry (20 WPM) was a great deal faster than multi-tap (8 WPM). They noted that actual performance of users did not match any of the models of expert entry.

---

[3] www. tegic.com

T9 is not the only dictionary-based disambiguation system available in commercial products. Other examples include eziText, developed by Zi Corp2 and Motorola's iTap. Unfortunately, no formal evaluations of these methods have been published.

As an alternative to dictionary-based disambiguation, Eatoni Ergonomics (www.eatoni.com) developed LetterWise. This method also uses linguistic knowledge to disambiguate key presses, but is not dictionary based. LetterWise uses a database of letter-prefix probabilities to disambiguate key presses (MacKenzie, Kober, Smith, Jones, & Skepner, 2001). Similar to T9, a list of predictions, ordered by probability, is generated after each non-space keystroke. These predictions are made on a letter-by-letter basis rather than for an entire word. Therefore, new keystrokes do not change the result of previous keystrokes. LetterWise also allows the input of non-dictionary words without switching input modes.

In their (MacKenzie et al., 2001) user study comparing LetterWise to multi-tap, the average text entry rate was 7.3 and 7.2 WPM for LetterWise and multi-tap, respectively. With training, entry speeds were 21 WPM for LetterWise and 15.5 WPM for multi-tap. Unfortunately, they did not compare LetterWise and T9.

2.4.3 Soft Keyboards

A soft or virtual keyboard is an image of a keyboard on a touch-sensitive screen or surface. Users enter text by directly tapping on the keyboard with a stylus or finger. Soft keyboard text entry, however, is a two-handed, eyes focused activity (MacKenzie & Zhang, 2001). The advantages of soft keyboards include their simplicity and that they are displayed on demand (MacKenzie & Soukoreff, 2002b). Additionally, they are easy to implement and change through software.

One major consideration for soft keyboards is the layout. In contrast to physical keyboards, a standard layout has been not established for soft keyboards. Comparative evaluations have found that participants with no experience using soft keyboards (but who do have experience using desktop computers), can achieve substantially higher initial text entry rates with a QWERTY layout than with alternative layouts (MacKenzie, Zhang, & Soukoreff, 1999). Consequently, the QWERTY layout is most frequently used in commercial products (such as information kiosks), where soft keyboards are used for minimal text entry by average users. The rationale for this is that the QWERTY layout can provide immediate usability by benefiting from skill transfer from desktop computing (Isokoski, 1999).

MacKenzie & Zhang (2001) suggest that experienced desktop users benefit from skill transfer when switching to a QWERTY soft keyboard. If the layout is unfamiliar, no skill transfer takes place. This suggests that visual scan time is a dominant component in determining text entry rates for novice users of alternative soft keyboards (Sirisena, 2002).

From this, it is reasonable to expect that the ordering of the Roman alphabet would be as familiar as the QWERTY layout to novice users. However, this has not been proven to be the case. In fact, the immediate usability of the alphabetical layout has consistently been less than expected, despite the believed familiarity of the layout (MacKenzie et al., 1999; Norman & Fisher, 1982), a result typically attributed to the alphabetical discontinuity caused by row breaks (Norman & Fisher, 1982; Zhai & Smith, 2000), given users' familiarity with the alphabet as a continuous ordered sequence and not as a set of discretely ordered rows (Sirisena, 2002). Therefore, users' must learn the breakpoints of the keyboard (determined by the number of rows) before realizing the benefits of the skill transfer.

Modeling user performance with soft keyboards has received significant attention in the past years (Isokoski, 2004). For example, Soukoreff and MacKenzie (1995) presented a model to predict expert and novice performance using a stylus to tap on a soft keyboard. Their model had three components: Fitts' Law (to predict movement time), Hick-Hyman Law (to predict visual scanning time), and linguistic tables of English digrams. The model predicted text entry rates of 30.1 WPM with a QWERTY and 8.9 WPM with an alphabetic layout.

As a follow up to Soukoreff and MacKenzie's (1995) predictive model for novice users, Sears, Jacko, Chu, & Moro (2001) studied the same six soft keyboards as those used by MacKenzie et al. (1999). Their study, however, was designed to estimate visual search times for soft keyboards independent of movement time. The study findings indicated that the patterns of visual search time are not consistent with the Hick-Hyman Law (Lewis et al., 2008). The modeling of novice performance remains a gray area (Isokoski, 2004).

There have been significant efforts to improve the design of soft keyboards. These improvements have generally resulted in alternative layouts, changing the positions of keys to minimize the time and effort for stylus and/or finger movement. Given that the use of a soft keyboard is based on tapping with one hand, typing speeds can be estimated by Fitts' Law analysis (Silfverberg, 2007). Some of the proposed layouts, such as POBox (Masui, 1998), OPTI (MacKenzie & Zhang, 1999), Metropolis (Zhai, Hunter, & Smith, 2000), ATOMIK (Zhai et al., 2002) and a few others are presented herein.

Masui (1998) developed an intelligent keyboard called POBox (Pen-Operation Based On eXample). The keyboard layout is static, but on each key press a fluctuating menu, listing the most probable completions for the current word, appears (see Figure 2-4). POBox uses a dictionary search and a digram language model to predict the next most likely words.

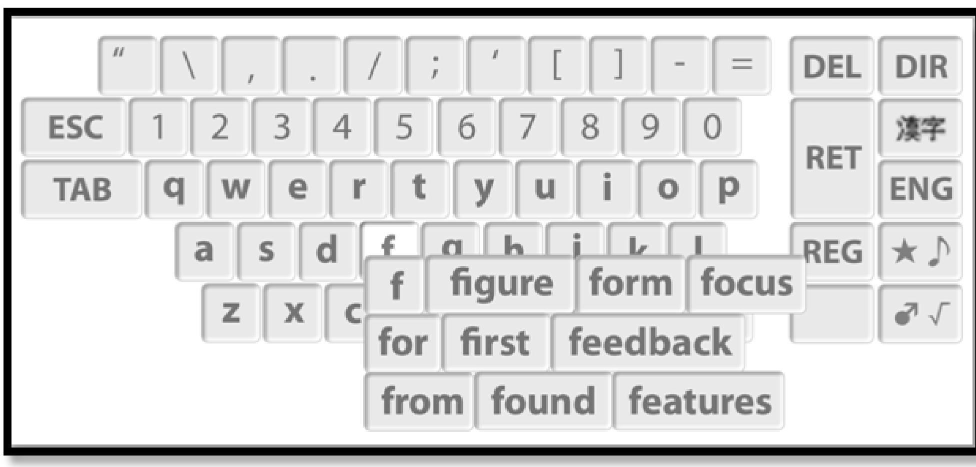


Figure 2-4 POBox Keyboard

Using Soukoreff and MacKenzie's (1995) model, MacKenzie and Zhang (1999) developed OPTI, a digram-optimized keyboard arrangement. In OPTI, the ten most likely characters are in the center of the keyboard, and the ten most frequent digraphs are on the top ten keys. The layout includes four evenly distributed space keys, while the remaining letters are on the remaining keys (see Figure 2-5). Current predictions of the optimized keyboard stand at about 42 WPM (MacKenzie & Soukoreff, 2002b). In a longitudinal study, MacKenzie and Zhang tested OPTI against the QWERTY layout. Initial input speeds were 17 and 28 WPM for OPTI and QWERTY, respectively. At the final session, the respective rates were 44 and 40 WPM.

Figure 2-5 OPTI Keyboard

Lewis, Kennedy, and LaLomia (1999) also tried to optimize text entry for soft keyboards. They used Fitts' Law, English digrams, and path analysis to evaluate alternate layouts. Based on their network analysis, they developed a digram-based layout estimated to attain throughput 27% better than QWERTY. In a user study, initial text entry rates were 24 to 27 WPM for the QWERTY layout, 15.6 to 20.4 WPM for their alphabetic layout, and 12 to 14.4 WPM for their digram-based layout (Lewis et al., 2008; Lewis, LaLomia, & Kennedy, 1999).

Zhai, Hunter, and Smith (2000) developed a mathematically optimized keyboard, in which the keys are organized in a hexagonal grid (see Figure 2-6). The predicted expert text entry rate was around 43 WPM. Unfortunately, longitudinal evaluations for this layout are not available, so the attainable rates in practice are unknown (MacKenzie & Soukoreff, 2002b).

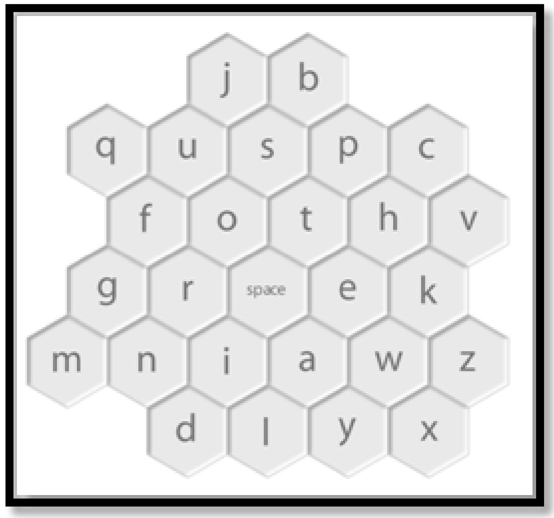


Figure 2-6 Metropolis Keyboard

Ward, Blackwell, and MacKay (2002) developed a novel data entry method called Dasher. With Dasher, users make continuous mouse movements across a stream of characters arranged by a predictive language model. A potential drawback of this approach is that the user has to continuously recognize dynamically rearranged letters (Zhai et al., 2005). Therefore, visual scan time may limit the performance of text entry with this method. Ward et al. (2002) reported text entry speeds of up 34 WPM from an initial study.

Zhai, Hunter, and Smith (2002) developed a layout, called ATOMIK. Given that for novice users the need to search for a target outweighs the movement time, ATOMIK introduced alphabetic ordering as a component of its development model to ease a user's search process. Consequently, ATOMIK is an optimized layout that is alphabetically tuned (see Figure 2-7).

Figure 2-7 ATOMIK Keyboard

In contrast to commercial layouts, most of the soft keyboards proposed in the literature only consider the entry of lowercase letters and spaces. For any keyboard to be usable in the real world, however, it must also provide means for performing functions and entering additional characters, such as symbols and punctuations.

2.4.4 Text Selection

A selection keyboard is an on screen keyboard that has a moveable selector (or cursor) over its keys. Using text selection (or few-key methods) on mobile devices is favorable as it reduces the need for larger screen areas to accept stylus or finger entry. Key-based selection techniques are possible with as few as two keys. In a two-key method, the user presses one key to scroll through a wheel of characters and uses the other key to select a character. Similarly, in a three-key text entry device, two keys are used to move the selector Left and Right, while the third key is used to select the desired character. With a five-key method, four directional keys move the selector over a matrix of characters and the fifth key is used for selection.

Selection techniques are prime candidates for text entry in mobile, wearable and assistive contexts. There are no commercial examples of the two or three-key methods in mobile computing. However, five-key text entry is widely used in consumer products such as television and gaming controllers. For example, both the Tivo DVR and Sony Xbox use a five-key method for text entry (see Figures 2-8 and 2-9). Typically, selection keyboards available in industry use either alphabetic or QWERTY Layouts.

A major advantage of text selection methods is that they are easily learned (Wobbrock, 2006) and can be operated with one hand. These text entry techniques require that the user keep track of the position of the selector over the matrix of letters. Therefore, this method requires high concentration and eyes-on operation. Furthermore, this type of text entry can be slow and inefficient (Sandness, Thorkildssen, Arvei, & Buverud, 2004). The following sections describe text entry inventions that use key-based text selection techniques.

Figure 2-8 Tivo DVR Uses Five Keys and an On-screen Keyboard for Text Entry

Figure 2-9 On-screen Keyboard and Gaming Controller for the Sony Xbox

2.4.4.1 Five-Key FOCL

Bellman and MacKenzie (1998) proposed an optimized five-key method called the Fluctuating Optimal Character Layout (FOCL). The FOCL strategy relies on letter-pair (digraph) probabilities to dynamically rearrange letters (based on the most recently selected letter) for the purpose of minimizing selector movement. Their FOCL keyboard used three rows to display the character set and included a fixed space, spanning across the three rows, positioned to the left of the letters. The FOCL method supported only lowercase letters and the space, and used a *snap-to-home* cursor, where the cursor snaps back to a predefined home position after character selection. In their design, the home position was the top leftmost letter in the layout. Users pressed four directional keys to move the cursor over the on-screen keyboard and an Enter key for character selection. The initial layout (which was also the layout that would appear after entering a space) of the FOCL keyboard appears in Figure 2-10.

Figure 2-10 Five-Key FOCL Keyboard

To compare the FOCL strategy against a fixed layout, a five-key QWERTY keyboard was also developed. Eleven participants entered text (only lower case letters, words, and spaces) using each method. Error correction mechanisms were not available, so participants were asked to ignore mistakes.

Results showed no significant difference between methods in the average text entry rate (~10 WPM) or error rates. Although the study was longitudinal, Bellman and MacKenzie argued that participants did not have enough practice with the FOCL method to reach the crossover point, much less to attain their maximum text entry rates. Although the FOCL method greatly reduces the number of keystrokes required to enter a character, users need more visual search time to find the intended character (MacKenzie & Soukoreff, 2002b). Therefore, with more practice users should become more familiar with the dynamic layouts, minimizing visual search time. However, the arrangement of the letters (Figure 2-11), meant to reduce keystrokes to the more likely next character, also imposes an unnatural search pattern that may be very difficult to overcome, even with extensive practice.

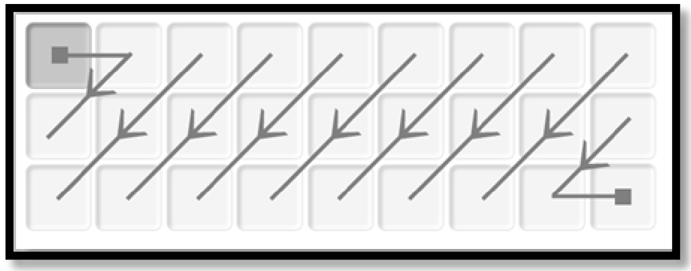


Figure 2-11 Letter Arrangement for the FOCL Keyboard

2.4.4.2 Three-key FOCL

MacKenzie (2002b) evaluated six text-selection techniques, each using only three keys for text entry and one row to display the character set. The techniques used the Left and Right arrow keys to move the key cursor over a linear arrangement of characters (letters and Space) and an Enter key to select the intended character. Four of the six methods evaluated employed alphabetic layouts. These methods differed only in the placement of the space key and the cursor mode. The two cursor modes evaluated were persistent and snap-to-home. A persistent cursor remains at the selected position following character selection. A snap-to-home cursor, as described previously, jumps back to the assigned key (the cursor home position) immediately after each entry.

The remaining two methods used Bellman and MacKenzie's (1998) FOCL strategy. For the FOCL methods, the space character was placed to the left and the cursor mode was snap-to-home, with the home position assigned to the space character. The FOCL methods differed only in their letter arrangements. In one method ("FOCL Level 1"), the letters were in a digram-optimized order, with the next most likely letter (based on the preceding entry) positioned after SPACE. For the other FOCL method ("FOCL

Level 2"), the order was determined by the preceding two characters (trigram optimization).

Additionally, all six methods allowed typematic movement for accelerated input. Typematic (also known as auto repeat) allows users to produce a continuous stream of virtual key presses by pressing and holding the key beyond a specified delay threshold, with virtual key presses generated at a fixed rate.

MacKenzie (2002b) used KSPC to evaluate the efficiency of each method. The two most promising methods were: 1. an alphabetic layout placing the Space to left of the letters and using a snap-to-home cursor and 2. FOCL Level 2. In an empirical evaluation, 10 participants entered text with these two methods (only lower case letters, words, and spaces). Participants ignored mistakes because the methods did not provide any error correction mechanism.

The results for entry rates, ~9-10 WPM, did not significantly differ, nor did error rates, ~2%. As in Bellman and MacKenzie's (1998) effort, the FOCL strategy was characterized as requiring high levels of concentration to navigate its constantly changing keyboard layout. Subjective measures indicated that participants preferred the alphabetic layout.

2.4.4.3 Multi-Ring, Tree-Based, and Binary methods for three-key text entry

Sandnes, Thorkildssen, Arvei, and Buverud (2004) developed four three-key text entry strategies. User studies were conducted on the three theoretically best methods: Multi-ring, Tree-based, and Binary. In the Multi-ring method, characters were organized in groups and text was entered in two steps. The first step required that the user select a character group using the Left, Right and Select keys. In the second step, the user navigated within the group to select the desired character. With the Tree-based method, character entry required three steps. In the first step, users selected one of three groups that contained the desired character by pressing the key mapped to the desired group. Next, the user selected the subgroup, among three, that contained the desired character. Lastly, the desired character was selected by pressing the key associated with it. In the Binary method, users pressed one of the three keys to select groups of letters until acquiring the target letter, with the letters divided into two groups.

Twelve participants completed test tasks using each device. The test texts did not include numbers, punctuation, or capitalization. The results for entry rates were 2.3, 2.7 and 2.3 WPM for the Multi-ring, Tree-Based, and Binary methods, respectively. The observed mean KSPC were 6.33, 3.13, and 5.01 for the Multi-ring, Tree-Based, and Binary methods, respectively. Subjective measures indicated a preference for the Multi-ring method.

Chapter 3. Keyboard Design

This chapter introduces a new hybrid text entry technique. This method draws on designs discussed in Chapter 2; such as Bellman and MacKenzie's (1998) five-key FOCL strategy and MacKenzie's (2002b) three-key text entry techniques to guide the new hybrid design.

For this research, the design, interaction, and implementation issues are explored without association with any specific form factor. Specifically this work addresses five-key input, but the techniques are extensible to other interaction controls such as five-way navigation, trackball, and joystick. As discussed in section 2.3, optimization strategies are generally language dependent. The language considered in this effort is English.

3.1 Design Goals

The motivation for this effort was to design and evaluate text input strategies suited to small, input-limited, handheld devices, satisfying the following objectives for the purpose of optimizing for minimal cursor movement and maximum usability:

1. The character set must appear on an on-screen keyboard and must not occupy more than two lines on the display. This model is appropriate for displays that are more wide than tall, such as those found on mobile phones, two-way radios and handheld gaming consoles.
2. Support one-handed operability.
3. Use only four keys to navigate (Up, Down, Left, and Right) through the character set and one for selection.
4. Reduce keystrokes as much as possible.

5. Provide an extended character set that, at minimum, includes the following:
    a. Numbers: 0-9
    b. Punctuations: period (.), comma (,), question mark (?), exclamation mark (!), single quote (‘), double quote (“), and the colon (:)
    c. Symbols: underscore (_), dash (-), forward slash (/), and the at sign (@)
6. Include a subset of edit and modifier keys, such as Space, Backspace, Shift, and Enter.
7. Develop a hybrid layout that combines fixed alphabetic and fluctuating predictive keys. To evaluate performance gains in using a hybrid keyboard, it is also necessary to develop an Alphabetical (or fixed) and a Predictive (or FOCL) keyboard.

The overall goal of this design effort was to improve five-key text entry speeds on constrained devices.

3.2 KSPC- As an Analysis Tool

As explained in section 2.2.2, calculating KSPC requires the use of keystroke data and a language model built using a representative body of text. This work uses a “common” English language model (Mayzner & Tresselt, 1965; Soukoreff & MacKenzie, 1995).

In 1965, Mayzner and Tresselt documented a table giving the 26 x 26 letter-pair frequencies in common English. Their work included the sampling of 20,000 words from various sources. However, their work did not include the space character. Because the space is the most common character (and e-space the most common digraph) in text entry, Soukoreff and MacKenzie (1995) extended this table to include the space. Their table contains 729 (27 x 27) entries and the associated frequencies for each pair of characters (the number of times the second letter occurred immediately following the first, converted to a probability by dividing the frequency by the number of characters (107,199) in the sample). For example, the five most frequent digrams (using "_" for the space) and their probabilities are "e_" (4.57%), "_t" (3.65%), "th" (3.52%), "he" (2.94%) and "d_" (2.45%).

Other researchers have constructed digraph tables using different sources and sampling more words than Mayzner and Tresselt (1965). For example, Zhai, Hunter, and Smith (2002) constructed two letter-pair tables, one based on on-line chat logs and the other based on a corpus of several newspapers. They noted that the differences between their tables and that of Mayzner and Tresselt were minute, and consequently decided, for the promotion of consistency in the literature, to use the Mayzner and Tresselt table in their continuing work (Zhai, Kristensson, & Smith, 2005) – a decision respected herein.

To compute an estimate of the average KSPC for any five-key design given a table of the letter pair probabilities and a table detailing the number of keystrokes needed to navigate from any character to the next, use:

$$\overline{KSPC} = \sum_{i \in N} \sum_{j \in N} (p_{ij} \times d_{ij})$$

where i is the first character, j is the second character of the digram, p_{ij} is the probability of occurrence for the letter pair, d_{ij} is the smallest number of keystrokes needed to get from i to j (plus one for the keystroke required to make the selection), and N is the character set.

3.3 Designing Five-Key Techniques

There are numerous factors that influence the design of on-screen keyboards. Some of these factors tend to interact, so it is important to consider the associated tradeoffs.

3.3.1 The Character Set

The characters occupy only two rows on the display. Any of the three keyboards display, at minimum, 46 characters (26 letters, 10 numbers, 6 punctuations, and 4 special characters). Numbers, punctuations, and symbols are in logical groups. The top row is mostly letters and the bottom row mostly the extended character set.

3.3.2 Fixed, Predictive and Hybrid Keyboards

Because a QWERTY layout would require three rows of letters, the fixed keyboard in these studies used an alphabetic layout. The main advantage of an Alphabetical keyboard is that users can easily become familiar with the layout, reducing visual search time to zero for the expert user. Unfortunately, this layout requires a greater number of keystrokes for entering text.

The Predictive keyboard used Bellman and MacKenzie's (1998) Fluctuating Optimal Character Layout (FOCL) strategy, see section 2.4.4.1. The main advantage of the FOCL strategy is that it significantly reduces KSPC. However, a fluctuating keyboard is more difficult to learn because the user must work with 27 different layouts, greatly increasing the demand of visual search.

The main purpose of a hybrid keyboard equipped with a full set of fixed characters and a limited set of dynamic keys was to gain the advantages associated with alphabetically ordered and pure FOCL keyboards. The resulting hybrid keyboard could allow users to take advantage of the reduced KSPC in the dynamic portion of the layout while allowing easy access to a fixed alphabetic layout when the desired next character is not one of the most likely next characters.

3.3.3 Wraparound Cursor

Wraparound can be enabled or disabled. With wraparound enabled, the cursor will not stop at the left or right end of a row (as it would with wraparound disabled), but will wrap around to the other side.

3.3.4 Cursor Mode

Cursor modes can be persistent or snap-to-home. A persistent cursor remains at the selected position following character selection. A snap-to-home cursor jumps back to the assigned key (the cursor home position) immediately after each entry. Therefore, evaluating snap-to-home cursor mode also requires assessing the best home position.

3.3.5 Cursor Home Position

The cursor home position is the location assigned to the cursor in a snap-to-home mode. As in Bellman and MacKenzie (1998), this work evaluated a subset of possible cursor home positions to investigate this variable. In English, the probability of the space character is 18% (Soukoreff & MacKenzie, 1995). Given its prominence in typing, it is common to grant the space special treatment (Bellman & MacKenzie, 1998). For this reason and because previous studies have found the approach promising (MacKenzie, 2002b; Wobbrock, Myers, & Rothrock, 2006), the space was the cursor home position for all layouts.

3.3.6 Space Location

In general, the designs placed the space key in a location that would enhance natural left-to-right scanning of the letters in the layout. The specific location differed among the layouts.

3.3.7 Error Correction

Errors and error handling are important dimensions to consider when designing keyboards. As a consequence, all keyboard designs in these studies included a backspace key above the keyboard (in line with the user output). Thus, adding the backspace key did not increase KSPC for any layout and it was possible to acquire the backspace in no more than three keystrokes.

3.4 The Keyboards

3.4.1 Alphabetical Keyboard

The fixed keyboard used an alphabetic layout. To evaluate the multiple designs for an alphabetic keyboard, KSPC was calculated for the 16 possible layouts given the design factors explored in this work. As detailed in Table 3-1, the KSPC for the alphabetic keyboards varied from 13.43 to 6.53. Figures 3-1and 3-2 depict main effects and interactions for various key design variables.

The main effect analyses showed that enabling wraparound dramatically reduces KSPC, placing the space key in the center of the layout significantly reduces KSPC, and using a snap-to-home cursor slightly reduces KSPC. Analysis with Minitab's "Response Optimizer" tool confirmed that the optimal alphabetic keyboard had wraparound, a centered space, and a snap-to-home cursor. The only interaction was between wraparound and space location. As is clear from Figure 3-2, there is very high variance in KSPC without wraparound. Enabling wraparound reduces KSPC sensitivity to the other factors, and always results in substantially lower KSPC. Thus, the Alphabetical keyboard design employed wraparound and a snap-to-home cursor. Although a center space location would have KSPC equal to 6.53, the space was placed in the top leftmost position to enhance natural left-to-right scanning of the characters, which increased KSPC only slightly to 6.86 (a 5% increase – see Figure 3-3).

Design	*Cursor Wraparound*	*Cursor Mode*	*Cursor Home Position*	*Space Location*	*KSPC*
1	Yes	Snap-to-home	Space	Center	6.53
2	No	Snap-to-home	Space	Center	6.53
3	Yes	Snap-to-home	Space	Top Leftmost	6.86
4	Yes	Snap-to-home	Space	Top Right	6.86
5	Yes	Snap-to-home	Space	Bottom Leftmost	7.23
6	Yes	Persistent	Not Applicable	Top Leftmost	7.89
7	Yes	Persistent	Not Applicable	Top Right	7.89
8	Yes	Persistent	Not Applicable	Bottom Leftmost	7.96
9	Yes	Persistent	Not Applicable	Center	8.16
10	No	Persistent	Not Applicable	Center	9.22
11	No	Snap-to-home	Space	Bottom Leftmost	10.54
12	No	Snap-to-home	Space	Top Leftmost	10.54
13	No	Persistent	Not Applicable	Top Leftmost	10.88
14	No	Persistent	Not Applicable	Bottom Leftmost	10.88
15	No	Persistent	Not Applicable	Top Right	11.56
16	No	Snap-to-home	Space	Top Right	13.43

Table 3-1 Sixteen Designs for the Alphabetical Keyboard

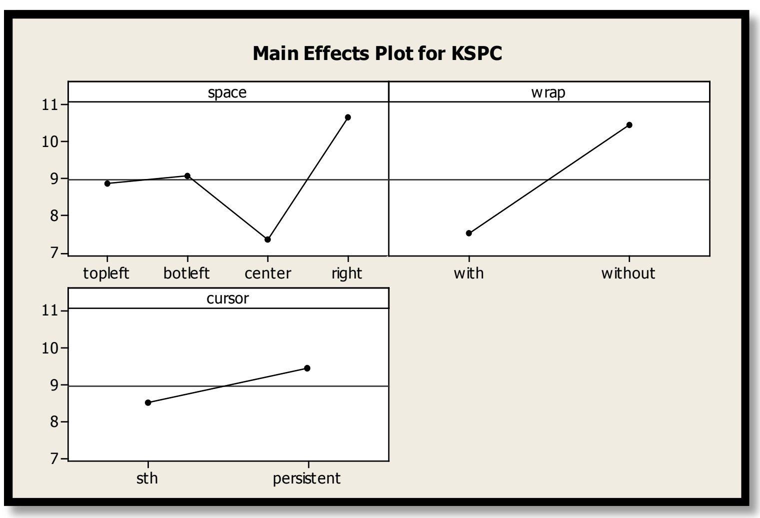


Figure 3-1 KSPC Main Effects Plot for the Alphabetical Keyboard

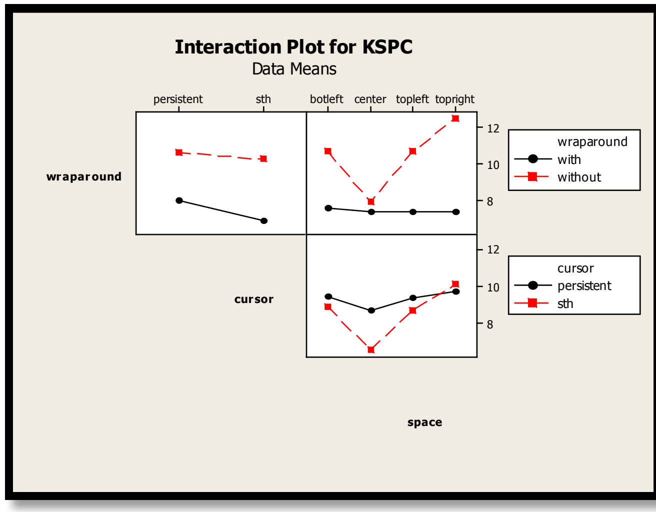


Figure 3-2 KSPC Interaction Plots for the Alphabetical Keyboard

Figure 3-3 Alphabetical Keyboard Layout

3.4.2 Predictive Keyboard

The Predictive keyboard used Bellman and MacKenzie's (1998) FOCL strategy, modeled to map closely to the three-key "FOCL Level 1" keyboard proposed by MacKenzie (2002b). The FOCL Level 1 layout places the space at the leftmost position and uses a snap-to-home cursor. This work (see Figure 3-4) used five-key navigation and a wraparound cursor, with an estimated KSPC of 4.52. This indicates that English text produced with the Predictive keyboard requires 34% fewer keystrokes per character than with the Alphabetical keyboard. The layout shown is the initial layout (also the layout that would appear after entering a space). All layouts (generated using the digram frequencies reported by Soukoreff and MacKenzie, 1995) appear in Table 3-2.

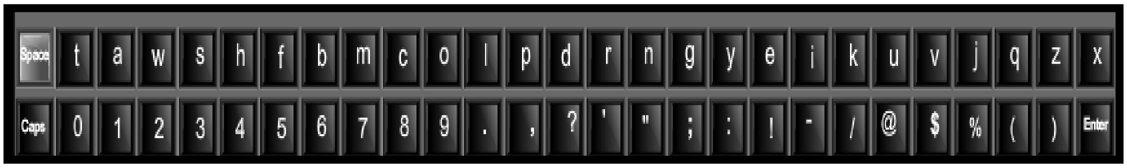


Figure 3-4 Predictive Keyboard Layout

Space	tawshfbmcolpdrngyeikuvjqzx
A	nrtsldiycvmkbgpufwxzhjaeoq
B	euoalrijsbytmvwcdfghknpqxz
C	haoektlriucysnqbdfgjmpvwxz
D	eioasruydlgvftnwbjmchkpqxz
E	rndaseltymcviwfpxgokhqbuzj
F	oireatflusymbcdghjknpqvwxz
G	ehoriaulsngtybmwcdfjkpqvxz
H	eaiotruyswlphmbcdfgjknqvxz
I	ntslcrdmgevokfpazxbuqihjwy
J	uoeiabcdfghjklmnpqrstvwxyz
K	einslyouampbcdfghjkqrtvwxz
L	elioadysutfkpmwvrgcbnhjqxz
M	eaoiupsbmyflntcdghjkqrvwxz
N	dgetosiyaclknujfhqvbrwxmpz
O	urnmtwolspvkdifbacyegzjxhq
P	eloariupsthywcfbdgjkmnqvxz
Q	uabcdefghijklmnopqrstvwxyz
R	eosiatydnulrmgkcvhfbpwjqxz
S	tehioauspklcywmnbqgdfjrvxz
T	heoaisrtuywlcbgndfmzjkpqvx
U	trsnlgcmeibpdafyozqxhjkuvw
V	eioayubcdfghjklmnpqrstvwxz
W	ahieonrsldtbukcfgjmpqvwxyz
X	tpaiceuobdfghjklmnqrsvwxyz
Y	oesipabltwghmucdfjknqrvxyz
Z	ezlasyiobcdfghjkmnpqrtuvwx

Table 3-2 Alphabetic Sequences Ordered by Likelihood to Follow each Letter

3.4.3 Hybrid Keyboard

According to Bellman and MacKenzie (1998), the FOCL benefit of reducing keystrokes only occurs when the target letter is close to the cursor. In the proposed Hybrid keyboard, the layout includes a fixed alphabetic row and a small number of fluctuating positions centered within the second row (see Figure 3-5).

The appropriate number, *n*, of fluctuating positions was determined by calculating the cumulative sum of probabilities associated with *n* positions, where *n* ranged from 1 to 27. The goal was that the probability of the target letter appearing in a fluctuating position be greater than 0.80. To satisfy this objective, there should be at least seven fluctuating positions (when $n = 7$, the probability of the target letter appearing in one of the seven fluctuating positions is 0.81).

Consistent with the results of fixed keyboard analysis, the Hybrid layout used wraparound and a snap-to-home cursor. Given the expectation that people will use the fluctuating component more often than the fixed portion of the layout, the space was assigned to a position to the left of the fluctuating characters. As in the Alphabetical keyboard, this should enhance natural left-to-right scanning of the characters in the fluctuating part of the keyboard. For this keyboard, the estimated KSPC is 3.87, requiring 14% fewer keystrokes per character in comparison to the Predictive keyboard and about 44% fewer compared to the Alphabetical.

Figure 3-5 Hybrid Keyboard Layout

3.5 Summary

Three techniques for five-key text entry on mobile devices are introduced. The KSPC for the methods were 6.86 for Alphabetical, 4.52 for Predictive, and 3.87 for Hybrid, suggesting that the Hybrid layout has the potential to be the best of the three keyboards.

KSPC is a useful tool that allows the characterization and comparison of text entry techniques before development and evaluation. However, using KSPC is not a substitute for a user test of text entry using high fidelity prototypes. Rather, it is a screening tool for the early identification of weak text entry strategies, most appropriately used before the commitment of development and testing resources.

Chapter 4. Evaluation Tool

To support this research and to provide an interactive, extensible research platform to the scientific community, the author designed the evaluation tool described in this chapter.

4.1 System Description

This tool (Millet, Asfour, & Lewis, 2009) supports the development and evaluation of selection-based virtual keyboards. It takes as input xml files that define test phrases and keyboard layouts. A session control interface presents the evaluation environment to the test participant and allows the moderator to control specific keyboard attributes. A data log captures session data in a portable comma-separated value (csv) format. This log contains a list of all typing key events that occurred throughout the test session.

The source code is in the Action Script (AS) language, the coding language for Adobe Flash (Flash). Flash is a programming environment that supports complex vector graphics and animations, and is compatible with most operating systems, including Microsoft Windows, Apple OSX, and Linux. The compiled executable (.exe, also compatible with most operating systems) file has an embedded Adobe Flash Player, so there is no need to download a player to run the tool, which starts with a simple click on the main .exe file.

4.2 System Features

There are many conditions that can influence the design and evaluation of new or refined text entry methods. This tool provides numerous design features, allows for comparative evaluations, and automates the test protocol and data collection when designing and evaluating text entry methods that use virtual keyboards and indirect selection-based input. Notable features of the tool are flexible character layouts, use of an extended character set, automation of test protocols and data collection, and design enhancements that include error handling, capitalization, different cursor modes, and typematic keying.

4.2.1 Flexible Character Layouts

The tool supports any character layout, allowing for the evaluation of standard (QWERTY and Alphabetic) and nonstandard layouts, both static and dynamic. The layouts are stored in an xml file, and it is possible to change the layout as a function of the last character typed. This feature permits the implementation of predictive keyboards based on digraph frequencies (as in Bellman & MacKenzie, 1998; MacKenzie, 2002b; Millet, Asfour, & Lewis, 2008). The tool also supports the display of keyboard layouts across *n* rows, where *n* is a user-specified attribute. Furthermore, adjustments to the application window can reduce or increase the size of the on-screen keyboard.

4.2.2 Use of an Extended Character Set

Much of the previous research in selection-based virtual keyboards has used test texts composed only of words and spaces (Bellman & MacKenzie, 1998; MacKenzie, 2002b; Sandnes et al., 2004). Rather than restricting future studies to alphabetic characters and spaces, this tool allows researchers control over the inclusion or omission of alphabetic and numeric characters, punctuation, symbols, edit functions, and modifier functions.

4.2.3 Automation of Test Protocol and Data Collection

Typically, text entry experiments require participants to transcribe text using the input method of interest. To support transcription, this tool reads the test phrases from a file and presents them to the user for input. As participants enter text, the tool records a timestamp and key code for each keystroke, saving these in a csv file for follow up analyses.

4.2.4 Design Enhancements

Error correction, capitalization, typematic keying, and specification of cursor home positioning and cursor movement are supported. Researchers can enable or disable most of these features.

4.2.4.1 Error Handling

Errors and error handling are important dimensions to consider when designing and evaluating text entry methods. Previous studies of selection-based entry (Bellman & MacKenzie, 1998; MacKenzie, 2002b; Sandnes et al., 2004) have not included error correction (only measuring words per minute, sometimes with a correction for errors), but the only way to measure true text entry throughput is with correct words per minute (Lewis et al., 2008). Consequently, the tool provides a backspace key[4] to allow participants to correct input errors.

4.2.4.2 Capitalization

For some real-world data entry, it is important for users to be able to produce both upper- and lower-case letters. To support these tasks, the keyboard xml file permits the optional presentation of an on-screen Caps key.

---

[4] This key behaves as a backspace, but is labeled "delete" in the user interface to align with existing consumer mobile phone products.

4.2.4.3 Cursor Modes (Positioning and Movement)

Different cursor positioning and movement techniques can affect the ease and efficiency of selection-based text entry (Millet et al., 2008), so researchers can enable or disable these features. Cursor positioning can be persistent or snap-to-home. A persistent cursor remains on the selected position after character selection. A snap-to-home cursor jumps back to an assigned key (the user-specified cursor home position) immediately after each character entry. The tool optionally allows cursor wraparound within or between keyboard rows. By default, the cursor home position is set to snap-to-home and cursor wraparound is enabled.

4.2.4.4 Typematic Keying

When enabled, the typematic (auto repeat) feature allows users to continuously produce a character by pressing and holding its key beyond a specified delay threshold. By default, typematic keying has its repeat delay (initiation) set to 250 ms with a repeat rate (continuation) of 50 ms (as specified in Lewis, Potosnak, & Magyar, 1997).

4.2.5 Keyboard Layouts

The keyboard layouts to be tested are stored in plain text (xml) files, easily created and edited using Microsoft Notepad or an xml editor. The application supports only one key file at a time, but the file can store multiple keyboard layouts. This file also specifies the characters to display and the order in which they should appear. The tool's extended character set includes:

- All letters: a-z
- All numbers: 0-9

- Punctuations: period (.), comma (,), question mark (?), exclamation mark (!), single quote ('), double quote ("), semicolon (;) and the colon (:)
- Other: tilde (~), at sign (@), number sign(#), dollar sign ($), percent sign (%), caret (^), ampersand (&), asterisk (*), underscore (_), dash (-), forward slash (/), backward slash (\), plus sign (+), equal sign (=), round brackets (( )), curly brackets ({ }), square brackets ([ ]), and angle brackets (< >)
- Edit and modifier keys: Space, Caps, and Enter

See Figure 4-1 for a sample of xml script that generates a static alphabetic layout with Space and Caps keys.

XML Script

```
<key>
  <layout1>
    <order order="Space">Space|a|b|c|d|e|f|g|h|i|j|k|l|m|Caps|n|o|p|q|r|s|t|u|v|w|x|y|z>
    </order>
  </layout1>
</key>
```

Resulting Layout

Figure 4-1 Sample XML Script and Keyboard Layout

This tool can also implement dynamic layouts (e.g., the fluctuating layout of Bellman & MacKenzie, 1998), where the layout changes as a function of the most-recently entered character. To generate a dynamic keyboard with *n* characters, researchers must specify *n* order rows in the keyboard file, with each row defining the layout to use for each character.

4.2.6 Test Files

In a typical text entry evaluation, the experimental task is to input text using the text entry method(s) under study. Generally, there are two types of tasks: text creation and text copy (transcription). In a text creation task, participants memorize or generate the source text (MacKenzie & Soukoreff, 2002b). This differs greatly from a transcription task, where the participant reads the text to enter (as in Butts & Cockburn, 2002; Curran et al., 2006; Dunlop & Crossan, 2000; James & Reischel, 2001; MacKenzie & Zhang, 1999, 2001).

Although text creation more closely mimics real world usage, transcription tasks are more common in empirical evaluations, mainly because it is difficult to measure the effects of non-behavioral aspects (such as “pondering”) when participants compose text (MacKenzie & Soukoreff, 2002b). Also, in text creation tasks it is difficult to ascertain whether or not a participant committed an error because it is sometimes not possible to know what the participant intended to enter. Consequently, transcription is the favored approach when evaluating text entry in laboratory settings (Wobbrock, 2007). For these reasons, this tool allows the evaluation of text entry when performing transcription tasks.

To automate the test protocol, the test file contains all of the test phrases for a given participant and specifies the order of presentation of the keyboard layouts to use.

Upon launch, the application reads the xml file containing the text phrases. During execution, the tool presents these phrases to the participant for input, with the test phrases presented to the user one at a time in the specified order.

The test file uses an xml format for creation and editing with Microsoft Notepad or an xml editor. Figure 4-2 shows sample xml script for the display of one test phrase ("Hello World") with one keyboard layout (keyboard_type= 1, as defined in an associated key.xml file). The script specifies a two-row layout (keyboard_row= 2) with 14 keys in the upper row (keys_per_row= 14). The cursor home position is the character in the first position (keyboard_home_location= 1), which in this example maps to the Space character (see Figure 4-1).

Sample Test Text Script

```
<test>
  <set1>
    <keyboard_type>1</keyboard_type>
    <keyboard_row>2</keyboard_row>
    <keys_per_row>14</keys_per_row>
    <keyboard_home_location>1</keyboard_home_location>
        <phrase>Hello World</phrase>
  </set1>
</test>
```

Figure 4-2 Sample Script for a Single Test Phrase

4.3 User Interface

The user interface consists of a large text field at the top of the screen where the stimulus phrase appears, an output field displaying the characters typed by the participant, and the character set of the input method under evaluation. The delete key appears above the keyboard (in line with the user output).

The cursor position appears as a grey box around a white character (see Figure 4-3). The input keys control the cursor on the screen, allowing users to navigate the character set and the delete key. The input keys are the Up, Down, Left and Right arrow keys (used for navigation) and the Enter key (used for selection).

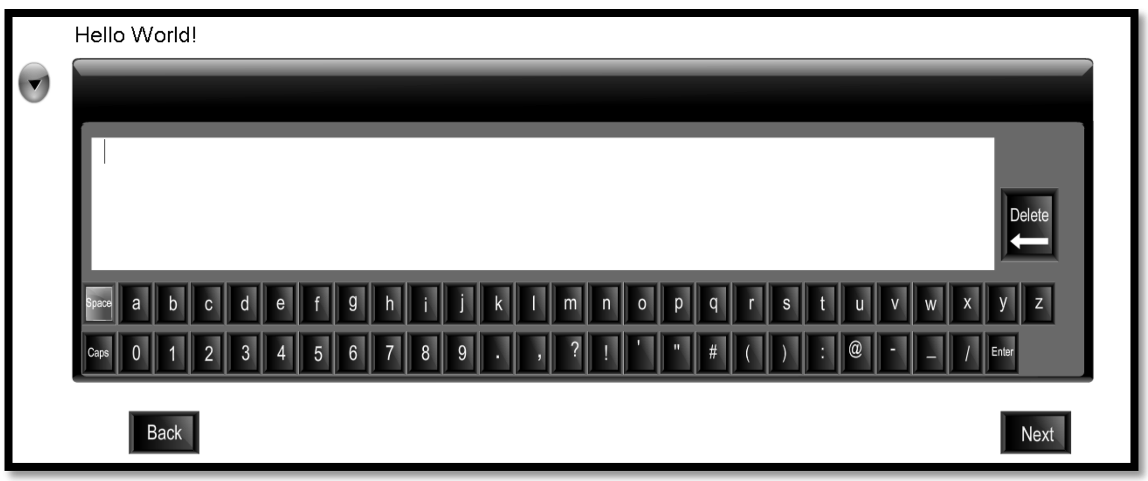


Figure 4-3 User Interface for Text Entry Evaluations

All other controls are accessed using a mouse. Typically, the experimenter (rather than the participants) will use these controls. Clicking the "Next" button signals the end of entering the presented test phrase and advances to the next test phrase. If the current test phrase is the last string in the test file, the "Next" button is disabled. Likewise, clicking the “Back” button returns to the previous test string. If the current string is the first phrase in the test file, then the “Back” button is disabled.

Pressing the down arrow button at the top left corner of the application displays the normally hidden main menu, which provides functions for uploading a specific test, starting the test, adjusting settings (via the “Options” item), ending the test (which automatically saves the associated data log file), and exiting the application (as shown in Figure 4-4).

Figure 4-4 Main Menu

As shown in Figure 4-5, the “Options” menu item allows the setup of several parameters that control important characteristics of the test keyboards: cursor home position, cursor wraparound, and typematic keying.

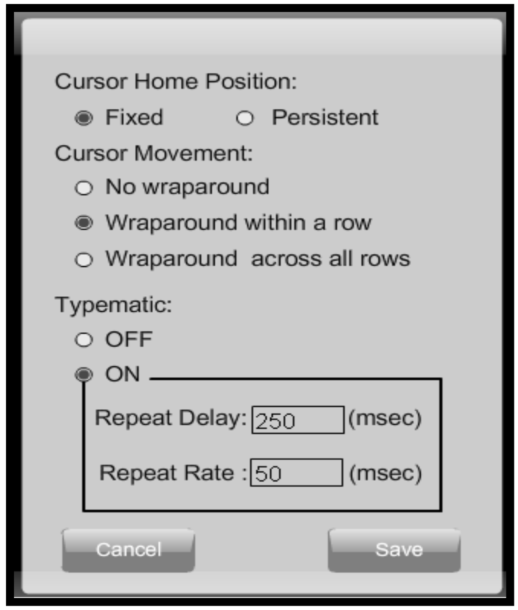


Figure 4-5 Options Menu

4.4 Data Collection

The tool's data logging of participant key presses (including arrow and selection keys) allows fine-grained analysis of sub-second events and low-level actions, measured to the nearest millisecond. Data collection begins upon initiating the test. Thereafter, the software records all participant interactions with the keyboards. The header of the data file captures the presented and the transcribed text, keyboard configurations, and the settings for each test session. Every subsequent line in the log file records the sequence of actions taken by the participant to generate the transcribed string, beginning with a timestamp followed by event codes and one or more identifiable handles. Specifically, for each key press, the tool collects the following information:

- The key pressed (e.g., Left, Right, Up, Down or Enter).
- User entered character, modifier or edit key.

- Time elapsed between key presses, or, for the last key, the elapsed time to clicking "Next".
- Indication if user employed typematic keying.

See Figure 4-6 for a sample portion of a data file.

Input:	Hello World!		
Output:	Hello World!		
keyboard Type:3			
Cursor Home Position:Fixed			
Cursor Movement:Wraparound with a row			
Typematic:ON			
Repeat Delay:250ms			
Repeat Rate:50ms			
Time (msec)	Keystroke	Notes	
2500	down		
3344	enter=Caps		
4047	right		
4453	right	typematic	
4531	right	typematic	
4609	right	typematic	
4688	right	typematic	
4875	right		
5031	right		
5156	right		
5859	enter=h		
6500	right		
6922	right	typematic	
7000	right	typematic	

Figure 4-6 Sample Data File

The tool gathers the data needed for the statistics and performance measures. To save the data, however, the test moderator must select "Done" from the main menu. The tool then saves the data as a csv file in the same location as the xml and .exe files, appending a date and timestamp to the file name to distinguish among multiple test sessions and prevent the accidental overwriting of data files. Researchers can retrieve data and use it for subsequent analysis, parsing and combining it as needed for the performance measures under investigation.

4.5 Summary

This chapter describes a tool for the evaluation of text entry methods for constrained devices. The main objective of this tool is to facilitate the design and implementation of prototypes of selection-based text entry methods, saving time and resources in the empirical evaluation process. The current version supports the flexible evaluation of selection-based virtual keyboard layouts for five-key text entry.

Chapter 5. Research Methods

A comparative evaluation was conducted to measure the performance of three five-key text entry techniques. The goal of the evaluation was to measure the performance of the Alphabetical (or ABC), Predictive (or FOCL), and Hybrid (or HYB) keyboards when entering different types of texts at three different levels of user training. All of the experiments were conducted with approval of the University's Institutional Review Board (IRB).

5.1 Test Environment

The experimental evaluations were conducted in a usability laboratory. Figure 5-1 shows the experimental workstation. The station consisted of a laptop, an external monitor, and a customized external keyboard used for cursor control (see Figure 5-2). Cursor control key mappings met the following criteria: (1) the Left/Right/Up/Down (navigation) keys formed an inverted-T shape and (2) the Enter key was beneath and to left of the navigation keys for easy acquisition by the thumb of the right hand (all participants were right-handed). Participants used their index, middle, and ring fingers for navigation and the thumb for selection. Participants remained seated during the experiment. The default height of the desk was its standard height of 26 inches, raised or lowered as required to accommodate varying participant heights. The monitor was directly in front of the user at an 18-inch viewing distance.

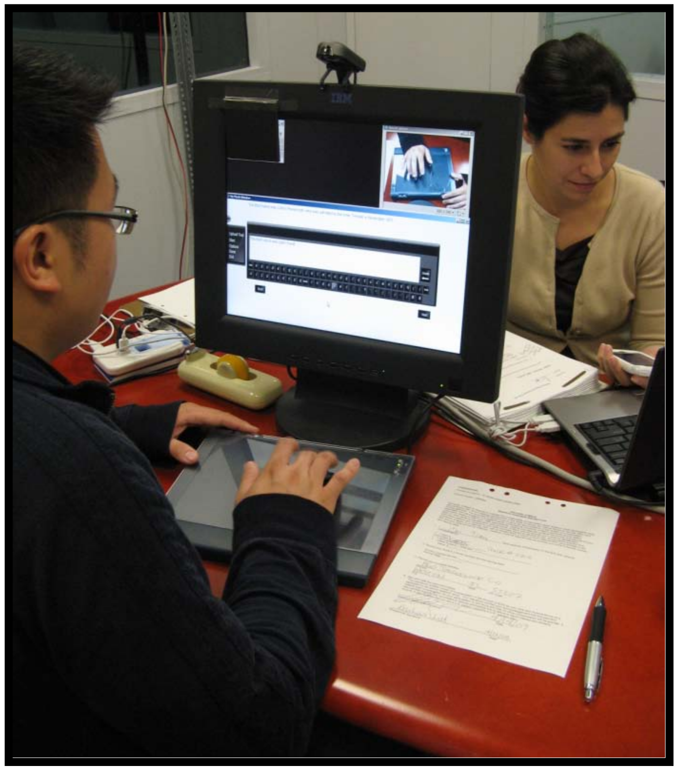


Figure 5-1 Experimental Setup for all Three Experiments

Figure 5-2 Customized External Keyboard Used for Cursor Control

5.2 Test Platform

The evaluation platform (or application), as described in Chapter 4, ran within a Microsoft Windows operating system and read the keyboard layouts and test phrases contained in two XML files. The application implemented the text entry techniques of interest. The user interface consisted of a large text field at the top of the screen where stimulus phrases appeared, an output field displaying the characters typed by the participant, and the character set of the input method under evaluation. The input keys controlled the cursor on the screen, allowing users to navigate the character set. Figure 5-3 shows a screen capture of the application. The xml code for the keyboard layouts and a sample test file are in Appendices D and E, respectively. Although the use of the keyboard and prototype application did not directly map to potential real world use, participants used the same apparatus for all input methods (in other words, all input methods received the same treatment).

Figure 5-3 Screen Capture of the User Interface

5.3 Data Collection

The evaluation platform automatically recorded all participant interactions with the keyboards. Data collection began with the first keystroke for each phrase and ended with the last keystroke. Specifically, for each test phrase the software collected both the presented and transcribed text with a time stamp and key code for each keystroke. Errors underwent manual analysis. For the purpose of this research, the level of analysis was the overall error rate (no breakdown of errors by type).

5.4 Experiment #1: Immediate Usability

The first stage of the evaluation tested the immediate usability of the three keyboards. As mentioned previously, immediate usability is the evaluation of a system after a user's initial exposure.

5.4.1 Participants

Participation was solicited by electronic postings. The participants were 24 adults, all right-handed, fluent in English, and with normal or corrected-to-normal vision. There was an equal mix of gender, age groups (<40 years old and ≥40 years old), and "texting" experience groups (Non-Expert and Expert). Modifying the criteria of Curran, Woods, and Riordan (2006), a non-expert was someone who sends fewer than 15 messages per week and an expert was someone who sends more. None of the participants had experience with any of the tested selection methods. The participants, in general, were highly educated, all owned mobile phones, and mostly reported average or greater proficiency with communication devices; see Tables 5-1 and 5-2. All participants could type at least 22 words per minute, see Figure 5-4, as rated by Typingtest.com, an online typing test developed by TypingMaster Finland, Inc. Participants received monetary compensation of $5 for their participation in this study.

Highest level of education achieved	*Frequencies*	*Percent*
High School Graduate	0	0
Vocational/ Technical Graduate	1	4.17
Some College	1	4.17
Bachelors Degree	13	54.17
Masters Degree	9	37.50
Doctoral Degree	0	0
Total	24	100

Table 5-1 Participants' Self Reported Highest Level of Education Achieved

Communication device proficiency	*Frequencies*	*Percent*
Novice	4	16.67
Average	14	58.33
Proficient	6	25.00
Total	24	100

Table 5-2 Participants' Self Reported Communication Device Proficiency

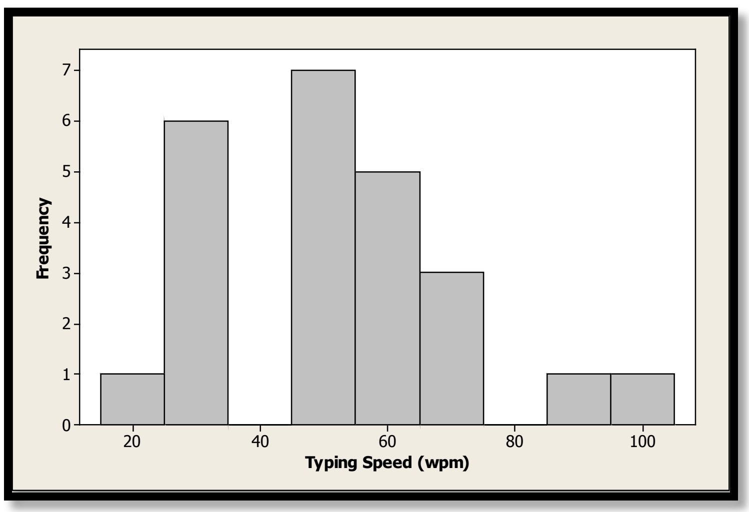


Figure 5-4 Participants' Typing Speed (in WPM)

5.4.2 Tasks

For the immediate usability task, participants typed the following phrase (43 characters including spaces):

the quick brown fox jumps over the lazy dog

This is a well-known (over-learned) phrase that contains each letter of the English alphabet.

5.4.3 Experimental Design

Each participant entered the test phrase once with each input method. The experiment was a mixed design with one within-subjects factor, keyboard layout (Alphabetical, Predictive, and Hybrid) and three between-group factors: gender (Male, Female), age (<40 years old, ≥40 years old), and experience (Non-Expert, Expert). The order of presentation of the keyboards was counterbalanced using a 3x3 Latin Square, with eight participants per order. The experiment addressed the following hypotheses:

- H_{01a}: For the initial reaction task, average entry speeds (in CWPM) for the Hybrid and Alphabetical conditions will be the same as those for the Predictive condition

  H_{A1a}: For the initial reaction task, average entry speeds (in CWPM) for the Hybrid and Alphabetical conditions will exceed those for the Predictive condition

- H_{01b}: For the initial reaction task, average subjective rating scores for the Hybrid and Alphabetical conditions will be the same as those for the Predictive condition

  H_{A1b}: For the initial reaction task, average subjective rating scores for the Hybrid and Alphabetical conditions will exceed those for the Predictive condition

5.4.4 Procedure

Participants completed a pre-test questionnaire soliciting demographic and mobile phone usage information (see Appendix A). Next, they received verbal instructions explaining the task and the goal of the experiment, including instruction to enter text as quickly and accurately as possible (see Appendix B). Participants then entered the test phrase once using each the three keyboard layouts without any practice trials, and could correct errors using the delete key. After entering the test phrase with each of the three keyboards, participants ranked the layouts in order of preference. Participants were tested one at a time, with participation lasting about 15 minutes.

5.5 Experiment #2: Novice Performance

The goal of this experiment is to assess user performance at the onset of learning in terms of text entry speed using each of three keyboard layouts and four types of text.

5.5.1 Participants

The participants in this experiment were the same as those in the first experiment. All participants received a payment of $15 as compensation for their time.

5.5.2 Tasks

The experimental task was the input of phrases of text. In this study, there were four types of text, designed to simulate realistic text entry when using internet-enabled devices. The text phrases encompassed characters across the sets of the evaluated keyboard layouts. This allowed evaluation of the relative ease in entering any character in the sets. Additionally, the text types chosen allowed for comparison of the results against the literature. The text types were:

- Words/Spaces: This task involved entering only lower case letters, words, and spaces. For consistency with prior experiments, the MacKenzie and Soukoreff phrase sets (MacKenzie and Soukoreff, 2003, modified to use only American English spellings and only lower case letters) were the source for the test texts (as in Clarkson et al., 2005; Gong & Tarasewich, 2005; Koltringer & Grechenig, 2004; Lyons, Starner, et al., 2004; Wigdor & Balakrishnan, 2004; Wobbrock et al., 2006). This text type represents the most tested task in the text entry literature

(Lewis et al., 2008). Random selection process from a base set of 500 phrases determined the specific phrases to use.

- Sentences: This task involved entering a few sentences consistent with writing a text or email message. As in prior work, the Brown Corpus was the source for a set of randomly selected sentences (Lewis, 1995). The Brown Corpus contains a selection of American English passages (500 samples, from books and magazines, distributed across 15 genres). All sentences selected ranged in length from 90 to 110 characters.
- Addresses: This task involved entering a few addresses. The purpose of the task was to assess the relative ease of entering numbers with the different keyboards. As in prior work, the study used addresses randomly selected from the Human Factors and Ergonomics Society member directory (Lewis, 1995).
- Web: This task involved entering relatively short URLs or email addresses, included to assess the relative ease of entering special characters with the different keyboards. This is the most rarely tested text type in current text entry research, mainly because most layouts do not have extended characters sets to allow the entry of these types of text. Therefore, it was necessary to create these texts for this study. However, the test texts of Sears and Zha (2003) used in their evaluation of QWERTY soft keyboards served as a model for this study's test texts for web addresses.

Sample phrases for each text type appear below, in Table 5-3. For all test phrases used in this experiment, see Appendix C.

Text Type	*Sample Phrase*
Words/Spaces	have a good weekend
Sentences	There are more than 7,500 known cultivars of apples resulting in a range of desired characteristics.
Addresses	8374 Maple Dr, Apt. 36–C, Baltimore, MD 21250
Web	www.travelocity.com/vaca23

Table 5-3 Sample Phrases for each Text Type

The phrases selected for the "Words/Spaces" and "Sentences" text types were representative of common English (verified by analysis of the letter frequencies of the phrases using a tool developed by Dr. I. Scott MacKenzie – available for download from http://www.yorku.ca/mack/phrasesets.zip), see Appendix C. There were nine test phrases for each text type – a total of 36 phrases. This was the source for the formation of three data sets, each containing three test phrases of each text type. Each data set contained a total of 12 phrases. The order of presentation of the phrases (and in turn the text types), within a phrase set, were randomized.

5.5.3 Experimental Design

The experiment used a 3x4 within-subject factorial design. The two factors were:

- Keyboard Layout: Alphabetical, Predictive, and Hybrid
- Text Type: Words/Spaces, Sentences, Addresses, and Web

The experimental design used a diagram-balanced Greco-Latin rectangle (Lewis, 1993b) to simultaneously counterbalance the presentation of keyboard layout, the phrase set, and the pairing of the keyboard layout and phrase set (see Figure 5-5, where letters represent the keyboard layout and numbers represent phrase set). The dependent measures were entry speed (CWPM), error rates (%), KSPC, typematic keying rate, and movement inefficiency[5]. As in prior work, the subjective evaluation used six ratings: ease of finding letters, ease of rapid input, ease of accurate input, ease of learning letter locations, ease of typing, and acceptability of keyboard layout (Lewis, 1995), explained in greater detail in the next section. After using the three layouts, participants ranked them. The participants' ratings and rankings served as secondary measures.

A2	C3	B1
B3	A1	C2
C1	B2	A3
B1	C3	A2
C2	A1	B3
A3	B2	C1

Figure 5-5 Pairs of Latin Squares

The experiment addresses the following hypotheses:

- H_{02a}: Novice average entry speeds when using the Hybrid and Alphabetical conditions will be the same as those when using the Predictive condition

[5] Typematic keying and movement inefficiency are text entry measures specific to selection-based methods. These measures are concerned with the amount and manner of selector movement. In this work, typematic keying rate is measured as the percentage of keystrokes produced by automatic key repeats, while movement inefficiency is the percentage of keystrokes exceeding the optimal.

H_{A2a}: Novice average entry speeds when using the Alphabetical and Hybrid conditions will exceed those when using the Predictive condition

- H_{02b}: Novice average subjective rating scores when using the Hybrid and Alphabetical conditions will be the same as those when using the Predictive condition

  H_{A2b}: Novice average subjective rating scores when using the Alphabetical and Hybrid conditions will exceed those when using the Predictive condition

- H_{03}: Novice average entry speeds using the Hybrid and Alphabetical conditions will be less than 5 CWPM (averaged across all text types)

  H_{A3}: Novice average entry speeds using the Hybrid and Alphabetical conditions will be equal or greater than 5 CWPM (averaged across all text types)

- H_{04a}: Average error rate (%) for the text types "Sentences", "Addresses" and "Web" will be same as that of text type "Words/Spaces"

  H_{A4a}: Average error rate (%) for the text types "Sentences", "Addresses" and "Web" will exceed that of text type "Words/Spaces"

- H_{04b}: Average entry speeds for text type "Words/Spaces" will be the same as those of text types "Sentences", "Addresses" and "Web"

  H_{A4b}: Average entry speeds for text type "Words/Spaces" will exceed those of text types "Sentences", "Addresses" and "Web"

- H_{05}: Novice mean error rate for the Predictive condition will be the same as those for the Hybrid and Alphabetical keyboard conditions

  H_{A5}: Novice mean error rates for the Predictive condition will exceed those for the Hybrid and Alphabetical conditions

5.5.4 Procedure

This experiment immediately followed experiment #1. At the start of the experiment, each participant received a brief tutorial on how to use the respective keyboard layout and time to familiarize themselves with the particular key mappings. Participants practiced entering text for two minutes using the same text as that used in the immediate usability task ("the quick brown fox …"). The training data was not a part of any subsequent analysis. After the practice period, participants began working with the first data set.

Overall, participants entered 12 phrases using the respective keyboard layout, with instruction to enter the phrases as quickly and as accurately as possible. If an error occurred, participants could make corrections by selecting the delete key.

As in Lewis' (1995) evaluation of writing and typing on small touch screens, when participants finished working with a keyboard layout, they rated the keyboard layout with Lewis' rating questionnaire (used with permission, presented in Figure 5-6). After participants completed the questionnaire, they could take a two minute break, and then repeated the process with the remaining keyboard layouts.

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

Figure 5-6 Lewis' Keyboard Layout Rating Form (1995)

Again, following the procedure in Lewis' (1995) evaluation of small touch screens, after completing the test tasks with all keyboard layouts, participants completed two additional forms. The first form asked participants to rate the importance of the various layouts' attributes (used with permission, presented in Figure 5-7), and the final form had participants rank the three keyboard layouts. Participants were tested one at a time, with participation lasting approximately 90 minutes, including three short breaks.

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

Figure 5-7 Lewis' Keyboard Layout Attribute Importance Form (1995)

5.6 Experiment #3: Expert Performance

To measure expert performance, it was necessary to conduct a longitudinal study to capture learning curves and estimate entry rates as users gained more and more experience with the three keyboard layouts.

5.6.1 Participants

All participants were recruited by electronic postings. The participants were 12 adults, all right-handed, fluent in English, and with normal or corrected-to-normal vision. There was an equal mix of gender, age groups (<40 years old and ≥40 years old), and "texting" experience groups (Non-Expert and Expert). Modifying the criteria of Curran, Woods, and Riordan (2006), a non-expert was someone who sends fewer than 15 messages per week and an expert was someone who sends more. Participants had no prior experience with any of the tested selection methods. The participants, in general, were highly educated, all owned mobile phones, and reported average or greater proficiency with communication devices, see Tables 5-4 and 5-5. Most participants could type at least 30 words per minute, as rated by Typingtest.com, an online typing test developed by TypingMaster Finland, Inc. Participants received monetary compensation of $200 for their participation in this study.

Highest level of education achieved	*Frequencies*	*Percent*
High School Graduate	1	8.33
Vocational/ Technical Graduate	1	8.33
Some College	2	16.67
Bachelors Degree	3	25.00
Masters Degree	3	25.00
Doctoral Degree	2	16.67
Total	12	100

Table 5-4 Participants' Self Reported Highest Level of Education Achieved

Communication device proficiency	*Frequencies*	*Percent*
Novice	0	0
Average	4	33.33
Proficient	8	66.67
Total	12	100

Table 5-5 Participants' Self Reported Communication Device Proficiency

5.6.2 Tasks

This evaluation used only one text type (Sentences). Participants entered three sentences (randomly selected from a source file of 180 sentences) for each keyboard in each session. As in Experiment #2, the source file of sentences was a random selection from the Brown Corpus. All sentences ranged in length from 90 to 110 characters. In addition, within each participant sentences were randomly selected without replacement to ensure that participants always entered unfamiliar phrases. To review all the test phrases for this experiment, see Appendix C.

5.6.3 Experimental Design

The experiment used a 3x20 within-subjects factorial design. The two factors were:

- Keyboard Layout: Alphabetical, Predictive, and Hybrid
- Session: 20 sessions

The design used a counterbalanced order of keyboard layouts, phrase sets, and phrases within each session to reduce confounding. Although all possible sequences of the conditions were not enumerated. The dependent variables and subjective measures were the same as in experiment 1 and 2. This experiment addressed the following hypotheses:

- H_{06a}: Average entry speeds for the Hybrid condition will be the same as those for the Alphabetical and Predictive conditions after 90 minutes of practice (~10 sessions)

  H_{A6a}: Average entry speeds for the Hybrid condition will exceed those for the Alphabetical and Predictive conditions after 90 minutes of practice (~10 sessions)

- H_{06b}: Average subjective rating scores for the Hybrid condition will be the same as those for the Alphabetical and Predictive conditions after 90 minutes of practice (~10 sessions)

  H_{A6b}: Average subjective rating scores for the Hybrid condition will exceed those for the Alphabetical and Predictive conditions after 90 minutes of practice (~10 sessions)

- H_{07}: Expert average entry speeds using the Hybrid condition will be less than 15 CWPM

  H_{A7}: Expert average entry speeds using the Hybrid condition will be equal or greater than 15 CWPM

- H_{08}: Expert average learning rates for the Hybrid condition will be the same as that of the Alphabetical and Predictive conditions.

  H_{A8}: Expert average learning rates for the Hybrid condition will be greater than that of the Alphabetical and Predictive conditions.

- H_{09}: Expert mean error rate for the Predictive condition will be the same as those for the Hybrid and Alphabetical keyboard conditions

  H_{A9}: Expert mean error rates for the Predictive condition will exceed those for the Hybrid and Alphabetical conditions

5.6.4 Procedure

Each participant completed 20 sessions, with sessions scheduled on weekdays (Monday through Friday) and separated by at least two hours, but not more than three days. As in the OPTI evaluation (MacKenzie and Zhang, 1999), this schedule simulated real system use while accommodating participants' work schedules.

Participants entered three sentences with each keyboard layout. Participants received instruction to enter the phrases as quickly and as accurately as possible. If an error occurred, participants could make corrections using the delete key. After participants completed sentence entry with a keyboard layout, they could take a two minute break, and completed the same final forms in the same order as in the novice evaluations, then repeated this process with the remaining keyboard layouts. Participation in the expert sessions did not exceed 45 minutes, including three short breaks.

Chapter 6. Results and Discussion

6.1 Approach to Analysis

This chapter presents and discusses the results of the experiments. For each experiment, the approach to analysis was similar, starting with justification of adjustments to the data and any removal of outliers, then with presentation of summary statistics and significance tests for the performance response variables and subjective measures. The performance variables were corrected words per minute (CWPM), keystrokes per character (KSPC), uncorrected error rate (UER), total error rate (TER), percentage of typematic events (TE), and movement inefficiency (MI). The subjective measures were keyboard ranks and keyboard ratings. The correlation and regression analyses have taken into account the repeated measures performed to test the significance of various factors of interest. The model assumptions and goodness of fit of the models were also evaluated. The data analyses were completed in Minitab 15, SPSS 17.0, S Plus 8.0, and Microsoft Excel 2007.

The models constructed relied on stepwise methods, specifically the use of the backward elimination method (unless specified otherwise), in which the initial model includes all predictor variables and their interactions. The contribution of each variable in explaining the variability of the response was quantified via p-values which were subsequently compared against a removal criterion ($p > 0.10$). If the predictor variable met the removal criterion, it was removed from the model and the model was re-estimated for the remaining predictors. The contribution of the remaining predictors was then reassessed. This chapter focuses on the re-estimated models. See Appendices F and G for the initial models for experiments 1 and 2, respectively.

Predominately, mixed designs ANOVA were used to determine the effects of factors. For tests of differences between the means of two samples, the t-test was used. Nonparametric tests, specifically Friedman's test with multiple comparisons (Lewis, 1993a), was used when the distribution of the data was non-normal.

To determine the effects of factors on binary outcomes, Logistic Regression incorporating repeated measures was used. Furthermore, the analysis of learning effects used regression analysis (fitting parameters for each participant and for session averages) and mixed effects modeling (to perform longitudinal analyses to understand the learning process and capture changes in time).

Normality was assessed graphically by examination of normal probability plots as well as computationally using the Anderson-Darling (AD) statistic and its associated p-value. When the data samples did not follow a normal distribution, transformations were applied in an attempt to normalize the data. If transformation was not successful, nonparametric methods were used. Homogeneity of variance was assessed graphically by examination of scatter plots of the residual versus predicted values. All graphics and test statistics used to test model assumptions appear in Appendices F, G, and H for experiments 1, 2 and 3, respectively.

The common criterion for statistical significance is a threshold of $p < 0.05$ using two-tailed p values (Lowry, 2005). This dissertation included many statistical tests corresponding to overlapping hypotheses. To control over the overall Type I error rate, α was set to 0.01 for all performance-based response variables. However, α was set to 0.05 for the subjective measures because they did not require the testing of overlapping hypotheses.

6.2 Experiment #1- Immediate Usability

6.2.1 Text Entry Speed

The overall results for average text entry speed (CWPM) were: Alphabetical M= 4.079 (SD= 1.056), Predictive M= 3.781 (SD= 0.996), and Hybrid M= 3.959 (SD= 1.096), as depicted in Figure 6-1. Table 6-1 provides descriptive statistics for CWPM by keyboard.

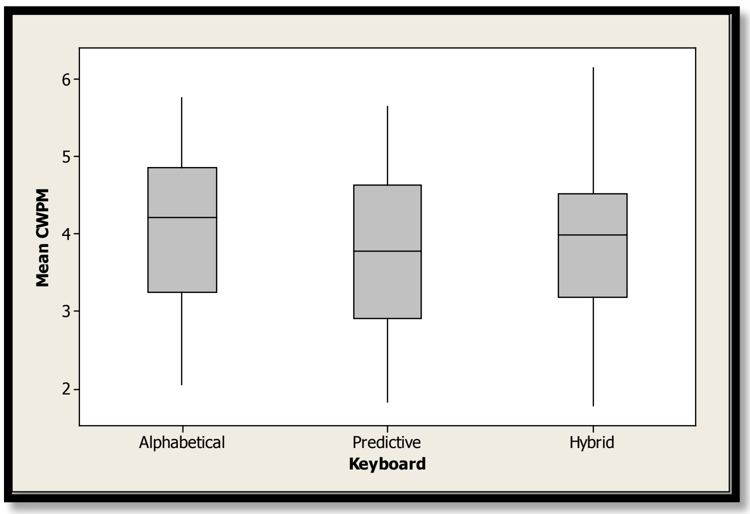


Figure 6-1 Mean CWPM for the Alphabetical, Predictive and Hybrid Keyboards

Variable	*Keyboard*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
CWPM	Alphabetical	24	4.079	1.056	2.050	3.248	4.215	4.857	5.770
	Predictive	24	3.781	0.996	1.820	2.903	3.780	4.628	5.660
	Hybrid	24	3.959	1.096	1.780	3.188	3.985	4.515	6.150

Table 6-1 Descriptive Statistics for CWPM for the Keyboards

There was a significant positive correlation between text entry speeds across the keyboards, specifically, between Hybrid and Alphabetical (r= .49, p= .014) and Hybrid and Predictive (r= .63, p< .001). This relationship indicates that if a participant performed well with one keyboard, then it is likely that participant also performed well with the other keyboards. A similar trend was observed between the Alphabetical and the Predictive keyboards (r= .35, p= .092), although it did not reach statistical significance. See Figure 6-2 for scatter plots depicting these relationships.

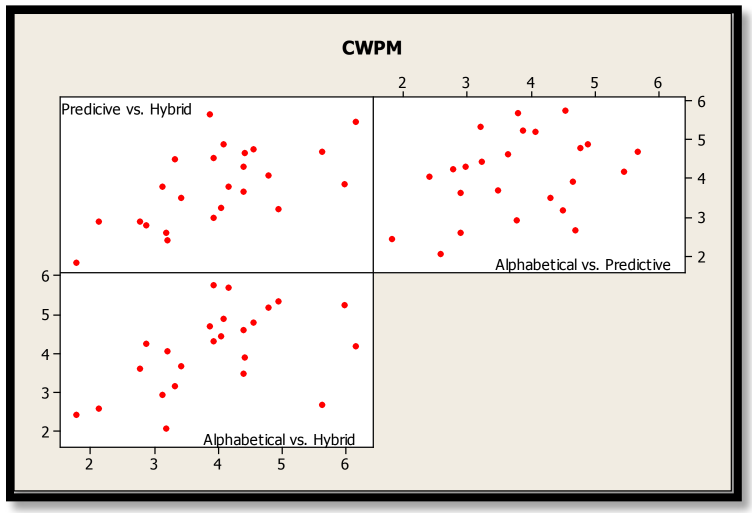


Figure 6-2 Scatter Plots of Keyboard Text Entry Speeds (CWPM)

To determine whether keyboard layout or other factors were associated with CWPM, repeated measures ANOVA was conducted. The final model for this design included keyboard, age, gender, and a gender by keyboard interaction. The results indicated that the mean CWPM was not associated with keyboards, $F(2,44)= 1.043$, $p= .361$ or with gender, $F(1,21)= 0.022$, $p= .883$. However, there was a statistically significant association between mean CWPM and age, $F(1,21)= 17.677$, $p< 0.001$ (Figure 6-3), with participants less than 40 years old faster than those more than 40 years old, on average. A gender by keyboard interaction effect was suggestive $F(2,44)= 3.085$, $p= .056$ (Figure 6-4), with pairwise comparisons indicating that males achieved greater speeds with the Alphabetical than with the Predictive keyboard (p = .01). The final model results, including the pairwise comparisons, appear in Table 6-2 and 6-3, respectively.

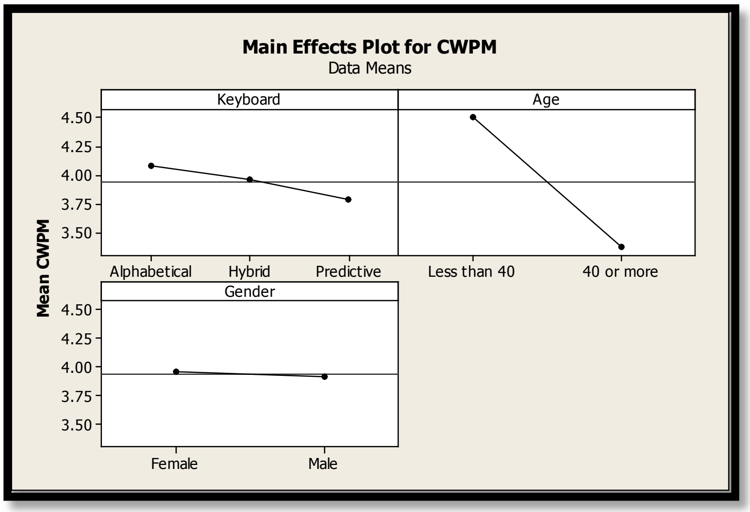


Figure 6-3 Main Effects Plots for CWPM by Keyboard, Age, & Gender

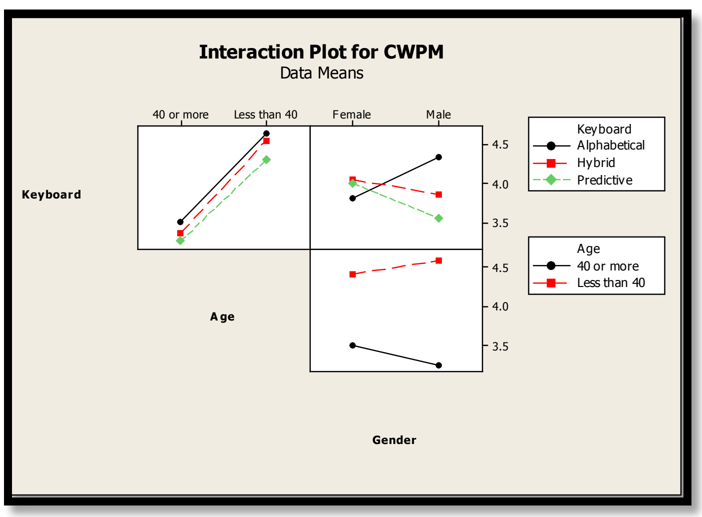


Figure 6-4 Interaction Plots for CWPM by Keyboard, Age, & Gender

Source	*Numerator df*	*Denominator df*	*F*	*Sig*
Intercept	1	21	859.116	.000
Age	1	21	17.677	.000
Gender	1	21	0.022	.883
Keyboard	2	44	1.043	.361
Gender*Keyboard	2	44	3.085	.056

Table 6-2 Tests of Fixed Effects for CWPM

Gender	*(I) Keyboard*	*(J) Keyboard*	*Mean Difference (I-J)*	*Std. Error*	*df*	*Sig*	*95% Confidence Interval for Difference*	
							Lower Bound	*Upper Bound*
Female	Alphabetical	Predictive	-.197	.293	44	.504	-.788	.393
		Hybrid	-.249	.293	44	.400	-.839	.341
	Predictive	Alphabetical	.197	.293	44	.504	-.393	.788
		Hybrid	-.052	.293	44	.861	-.642	.539
	Hybrid	Alphabetical	.249	.293	44	.400	-.341	.839
		Predictive	.052	.293	44	.861	-.539	.642
Male	Alphabetical	Predictive	.792	.293	44	.010	.202	1.382
		Hybrid	.490	.293	44	.102	-.101	1.080
	Predictive	Alphabetical	-.792	.293	44	.010	-1.382	-.202
		Hybrid	-.302	.293	44	.307	-.893	.288
	Hybrid	Alphabetical	-.490	.293	44	.102	-1.080	.101
		Predictive	.302	.293	44	.307	-.288	.893

Table 6-3 Gender by Keyboard Pairwise Comparisons for CWPM

6.2.2 Error Rates

6.2.2.1 Uncorrected Error Rate

The uncorrected error rate (UER) was the percentage of characters that remained in error in the transcribed string. Table 6-4 below provides descriptive statistics for UER. Initially a repeated measures ANOVA was conducted to determine whether keyboard layout was associated with UER. The model for this design included keyboard and age. The model results, presented in Table 6-5, indicate that the UER was not associated with keyboard layout ($F(2,46)= 1.429$, $p= .250$). The mean UER across age ($F(1,22)= 5.517$, $p= .028$) suggested that participants less than 40 years old made fewer errors, on average, than those who were more than 40 years old, but the finding was not significant (p> .01, see Figure 6-5). However, the test assumptions for the model could not be satisfied (see Appendix F). The UER data was significantly non-normal and zero inflated. Various transformations (such as square root and logarithmic) were explored, but none were successful.

Variable	*Keyboard*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
UER	Alphabetical	24	.012	.030	0	0	0	.015	.120
	Predictive	24	.018	.040	0	0	0	.020	.160
	Hybrid	24	.033	.062	0	0	0	.042	.210

Table 6-4 Descriptive Statistics for the Uncorrected Error Rate

Source	*Numerator df*	*Denominator df*	*F*	*Sig.*
Intercept	1	22	15.021	.001
Age	1	22	5.517	.028
Keyboard	2	46	1.429	.250

Table 6-5 Tests of Fixed Effects for Uncorrected Error Rate

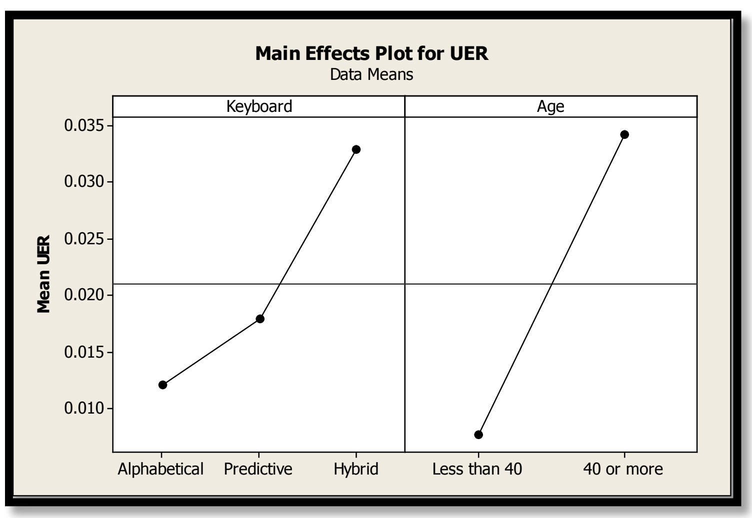


Figure 6-5 Main Effects Plots for Uncorrected Error Rate

Given that the UER data was non-normal and successful transformations could not be applied, Friedman's test was conducted. The test results indicate that the mean UER was not associated with Keyboard layouts, $X^2(2)= 2.68$, $p= .262$ (see Appendix F).

For an alternative analysis, UER data was coded to binary outcomes (assigned 0 for no errors occurring within a trial, and 1 for one or more errors occurring within a trial). A logistic regression incorporating repeated measures was then conducted to determine whether keyboard layout is associated with errors. The final model was determined via forward selection procedures, given that using the backward method was not possible due to insufficient degrees of freedom available to run the full factorial model. The final model for this design included keyboard and gender. Results indicated that keyboard layout ($X^2(2) = 1.339$, $p= .512$) was not associated with error, see Table 6-6. The proportion of errors across gender ($X^2(1) = 3.727$, $p= .054$) was suggestive, specifically that female participants were more likely to make errors than males.

Source	*Wald Chi-Square*	*df*	*Sig.*
Intercept	7.588	1	.006
Gender	3.727	1	.054
Keyboard	1.339	2	.512

Table 6-6 Logistic Regression- Tests of Model Effects for Uncorrected Error Rate

6.2.2.2 Total Error Rate

To explore the dynamics of error correction, errors committed before correction were investigated. The total error rate (TER) was calculated as the total number of error and corrective events divided by the total number of input events (both productive and corrective). Table 6-7 provides descriptive statistics for keyboard TER.

Variable	Keyboard	N	M	SD	Min	Q1	Mdn	Q3	Max
TER	Alphabetical	24	.094	.106	0	0	.065	.160	.440
	Predictive	24	.102	.094	0	0	.105	.182	.320
	Hybrid	24	.083	.791	0	0	.055	.160	.250

Table 6-7 Descriptive Statistics for Total Error Rate

A repeated measures ANOVA was conducted, to determine whether keyboard layout was associated with the TER. The model for this design was a full factorial including all factors investigated. The model results, presented in Table 6-8, indicated that the TER was not associated with any of the tested factors. Furthermore, none of the terms became significant after removing some of the interaction terms, so no further analysis was conducted for the TER.

Source	*Numerator df*	*Denominator df*	*F*	*Sig.*
Intercept	1	16.000	48.051	.000
Age	1	16.000	.300	.592
Gender	1	16.000	.845	.372
EXP	1	16.000	.451	.511
Keyboard	2	32	.240	.788
Age * Gender	1	16.000	.256	.620
Age * Experience	1	16.000	.146	.707
Age * Keyboard	2	32	.432	.653
Gender * Experience	1	16.000	.067	.799
Gender * Keyboard	2	32	1.339	.276
EXP * Keyboard	2	32	.308	.737
Age * Gender * Experience	1	16.000	.666	.426
Age * Gender * Keyboard	2	32	.362	.699
Age * Experience * Keyboard	2	32	.473	.627
Gender * Experience * Keyboard	2	32	.242	.787
Age * Gender * Experience * Keyboard	2	32	.622	.543

Table 6-8 Tests of Fixed Effects for Total Error Rate

6.2.3 Efficiency

6.2.3.1 KSPC

For the test phrase used in this study, the average computed minimal keystrokes per character (KSPC) for the tested layouts were 7.0 for Alphabetical, 4.9 for Predictive, and 4.7 for Hybrid. The overall results for average KSPC were: Alphabetical M= 11.94 (SD= 2.147), Predictive M= 8.373 (SD= 1.373), and Hybrid M= 7.633 (SD= 1.820), as depicted in Figure 6-6. Descriptive statistics for keyboard KSPC appear in Table 6-9.

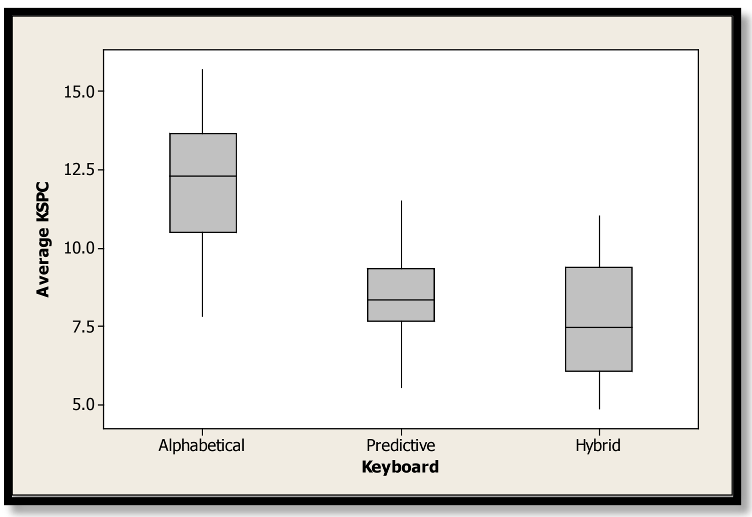


Figure 6-6 Average KSPC for the Alphabetical, Predictive, and Hybrid Keyboards

Variable	*Keyboard*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
KSPC	Alphabetical	24	11.940	2.147	7.840	10.510	12.315	13.663	15.690
	Predictive	24	8.373	1.373	5.560	7.668	8.345	9.338	11.510
	Hybrid	24	7.633	1.820	4.880	6.057	7.465	9.372	11.020

Table 6-9 Descriptive Statistics for KSPC

To determine whether keyboard layout, age, texting experience, or gender were associated with average KSPC, a repeated measures ANOVA was conducted. The final model for this design included keyboard, gender, age, experience, an age by experience interaction, and a gender by keyboard interaction. Model results, as detailed in Table 6-10, indicated a suggestive gender by keyboard interaction effect F(2,44)= 3.876, p= .028, see Figure 6-7), specifically that males may have an advantage over females when using the Alphabetical keyboard (p= .032, see Appendix F). The age by experience interaction effect was also suggestive (F(1,19)= 4.676, p= .044).

Source	*Numerator df*	*Denominator df*	*F*	*Sig.*
Intercept	1	19	1683.281	.000
Age	1	19	.184	.673
Gender	1	19	.038	.847
Keyboard	2	44	48.481	.000
Gender * Keyboard	2	44	3.876	.028
Experience	1	19	.307	.586
Age * Experience	1	19	4.676	.044

Table 6-10 Tests of Fixed Effects for KSPC

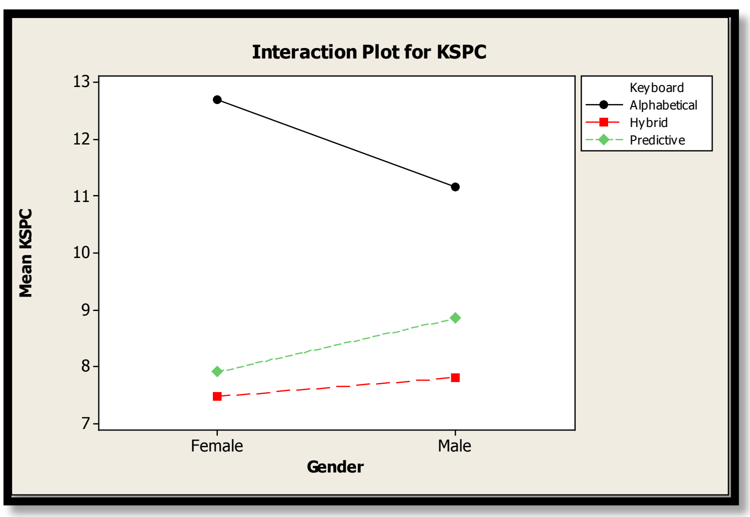


Figure 6-7 KSPC Interaction Plot for Keyboard by Gender

6.2.3.2 Typematic Events

In this work, a typematic event (TE) was any key sequence that began with a physical key press and extended to one or more virtual key presses through the use of auto-repeat. The length of the event was the sum of the virtual key presses. The data analyzed was the percent of TE as a function of total key presses (physical and virtual).

Typematic keying comprised a large percentage of the keystrokes in this study. Specifically, for the Alphabetical, Predictive and Hybrid keyboards, typematic events comprised 54%, 37%, and 27% of the observed keystrokes, respectively. Table 6-11 provides descriptive statistics for TE.

Variable	*Keyboard*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
TE	Alphabetical	24	.545	.210	0	.465	.595	.695	.760
	Predictive	24	.367	.209	0	.172	.415	.555	.650
	Hybrid	24	.269	.166	0	.130	.260	.415	.550

Table 6-11 Descriptive Statistics for Typematic Events

There was a significant positive relationship between the rates of use of typematic events between the Hybrid and Predictive keyboards ($r = .54$, $p = .006$), see Figure 6-8. This relationship indicated that if a participant employed typematic keying with the Hybrid keyboard, then it is likely that participant also employed typematic keying, with similar rates of use, with the Predictive keyboard, and vice versa.

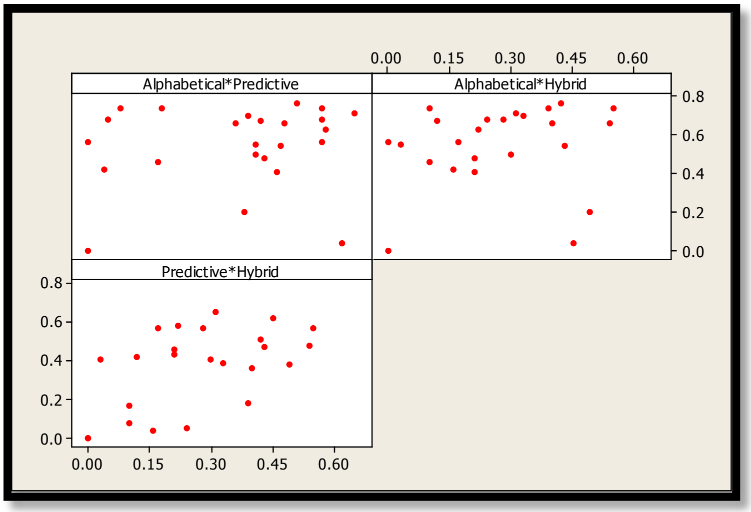


Figure 6-8 Scatter Plots of Keyboard Typematic Events

Initially a repeated measures ANOVA was conducted to determine whether keyboard layout was associated with the rate of use of typematic keying. The model for this design included only keyboard layout. The model results indicated that TE was associated with keyboard layout ($F(2,46)= 16.726$, $p < .001$). However, the test assumptions for the model could not be satisfied (see Appendix F) because the TE data was significantly non-normal. Various transformations were explored (such as inverse square root and logarithmic), but none were successful.

Given that the TE data was non-normal and successful transformations could not be applied, Friedman's test was conducted. The test results indicated that the TE rate of use was associated with keyboard layouts, $X^2(2)= 21.25$, $p< .001$ (see Appendix F). Specifically, TE rate of use for the Alphabetical keyboard differed significantly from that of Hybrid ($p< .0001$) and that of the Predictive ($p< .01$) keyboards. The greater typematic keying behavior for the Alphabetical keyboard was expected, given the inherently larger KSPC and greater predictability for letter locations. Likewise, less typematic keying behavior for the Predictive and Hybrid keyboards is reasonable as a consequence of their inherently lower KSPC.

In attempting an alternative analysis, the typematic event data was coded to binary outcomes (assigned 0 when typematic keying was not employed and 1 when typematic keying was used within a trial). There were, however, only five 0s out of 72 data points in the coded response, so logistic regression incorporating repeated measures could not be applied.

6.2.3.3 Movement Inefficiency

It is interesting to compare the path of observed selector movement to the minimal selector path for each of three keyboards studied. The minimal selector paths for the test phrases were calculated using a macro designed by the author. The observed path for the selector movement was the total keystrokes entered to generate the transcribed string, as recorded using the software tool. Subsequently, selector movement inefficiency (MI) is the percentage of keystrokes exceeding the optimal. For this experiment, participants, on average, moved the selector 71.3% more than optimal with the Alphabetical keyboard, 70.4% more with the Predictive, and 63.6% more than necessary with the Hybrid, as depicted in Figure 6-9. Table 6-12 provides descriptive statistics for Movement Inefficiency.

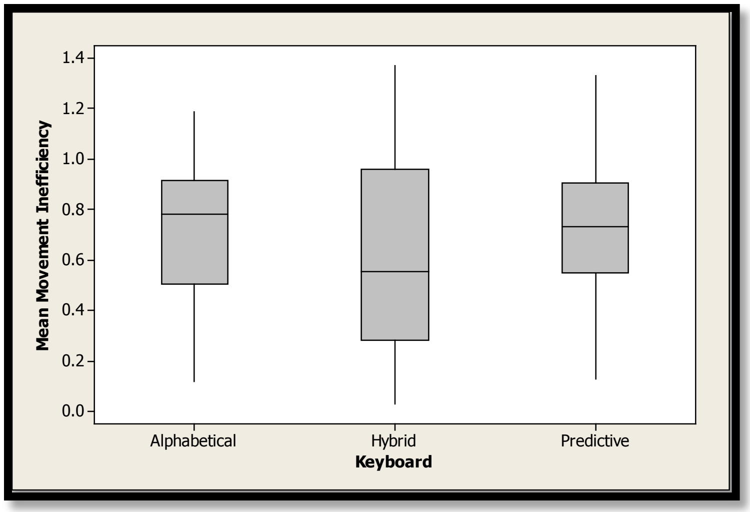


Figure 6-9 Mean Movement Inefficiency by Keyboard

Variable	*Keyboard*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
MI	Alphabetical	24	0.713	0.304	0.120	0.505	0.780	0.917	1.190
	Predictive	24	0.704	0.290	0.130	0.550	0.730	0.907	1.330
	Hybrid	24	0.636	0.416	0.030	0.280	0.555	0.962	1.370

Table 6-12 Descriptive Statistics for Movement Inefficiency

A repeated measures ANOVA was conducted. The final model for this design included keyboard, age, experience, and an age by experience interaction. The model results indicated that the mean MI was not associated with keyboards, $F(2,46)= .402$, $p= .671$. An age by experience interaction effect was suggestive ($F(1,20)= 7.026$, $p= .015$, see Table 6-13), but not significant.

Source	*Numerator df*	*Denominator df*	*F*	*Sig.*
Intercept	1	20	293.880	.000
Keyboard	2	46	.402	.671
Experience	1	20	.280	.603
Age	1	20	.596	.449
Experience * Age	1	20	7.026	.015

Table 6-13 Tests of Fixed Effects for Movement Inefficiency

6.2.4 Preference

The Hybrid keyboard received the most first-place votes (11/24), but an analysis of ranks indicated this finding was not statistically significant (Friedman test, X^2 (2) = 3.08, p= 0.214, see Figure 6-10). Participants who preferred the Hybrid keyboard stated that it required fewer keystrokes to acquire the target letter than the Alphabetical keyboard and had lower visual search demands than the Predictive keyboard.

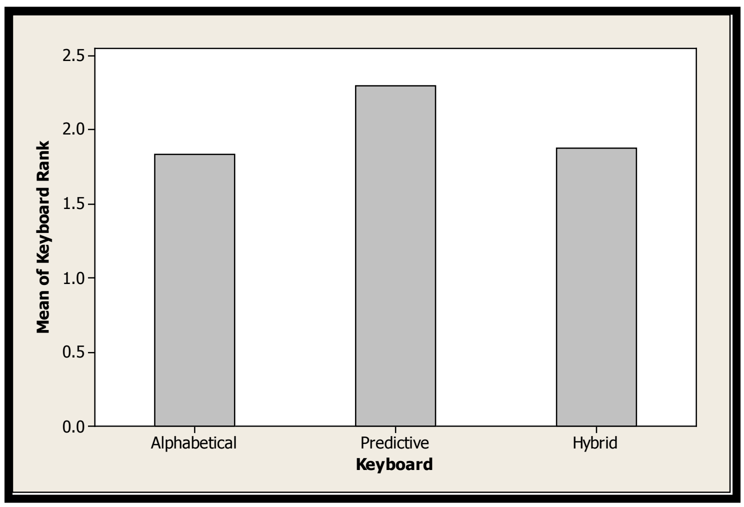


Figure 6-10 Keyboard Layout Ranks

6.2.5 Discussion

Experiment #1 investigated the immediate usability of three alternative keyboard layouts for supporting five-key text entry. There were significant performance differences as a function of age, with participants less than 40 years old superior to those greater than 40 years old. There were, however, no statistically significant differences (or even near-significant differences) in performance as a function of keyboard. Thus, the Predictive keyboard, despite its unusual layout, performed no worse than the Alphabetical and

Hybrid keyboards, suggesting that there may be no penalty associated with the immediate usability of this nonstandard layout for five-key text entry. Because the observed KSPC for all three keyboards were greater than the computed minimal values, it is likely that users can achieve even higher speeds with practice.

6.3 Experiment #2- Novice Performance

6.3.1 Text Entry Speed

The overall results for average text entry speed (CWPM) were: Alphabetical M= 4.980 (SD= 1.243), Predictive M= 5.030 (SD= 1.533), and Hybrid M= 5.087 (SD= 1.716), as shown in Table 6-14.

Variable	*Keyboard*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
CWPM	Alphabetical	288	4.980	1.243	1.885	4.148	5.039	5.776	8.106
	Predictive	288	5.030	1.533	1.513	3.967	4.802	6.016	10.152
	Hybrid	288	5.087	1.716	1.504	3.858	4.899	6.150	10.824

Table 6-14 Descriptive Statistics for CWPM by Keyboard

The overall results by text type for mean text entry speed were: Address M= 4.175 (SD= 0.923), Sentences M= 5.549 (SD= 1.387), Web M= 4.268 (SD= 1.079), and Words/Spaces M= 6.136 (SD= 1.542), as shown in Table 6-15.

Variable	*Text Type*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
CWPM	Address	216	4.175	0.923	1.625	3.688	4.230	4.840	6.232
	Sentences	216	5.549	1.387	2.166	4.770	5.635	6.346	8.941
	Web	216	4.268	1.079	1.504	3.503	4.373	5.018	6.952
	Words/Spaces	216	6.136	1.542	2.315	5.188	6.091	7.125	10.824

Table 6-15 Descriptive Statistics for CWPM by Text Type

Overall, there were significant positive relationships between text entry speeds across the keyboards. Pearson correlation coefficients were calculated for all pairs of the three keyboards (Alphabetical/Predictive ($r= .48$), Alphabetical/Hybrid ($r=.51$), and Predictive/Hybrid ($r= .44$)). Correlations were all significant ($p < .001$) and had similar trends between keyboard pairs. These results indicated that if a participant performed well with one keyboard, then it was likely that participant also performed well with the other keyboards. See Figure 6-11 for scatter plots depicting these relationships.

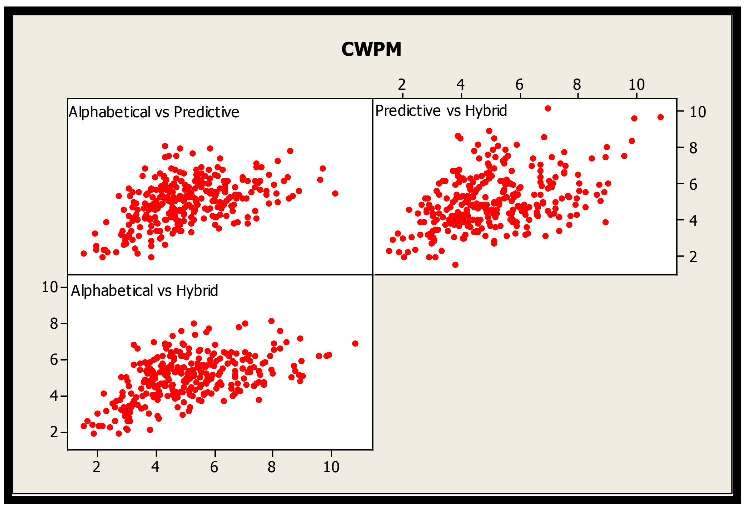


Figure 6-11 Scatter Plots of Keyboard Text Entry Speeds (CWPM)

One hypothesis for this experiment was that novice entry speeds would be greater than or equal to 5 CWPM (averaged across all text types). However, multiple one-sample t –tests provided no evidence that the observed entry speeds for any of the three keyboards was significantly greater than 5 CWPM (for Alphabetical $t= -0.28$, $p = .391$, Predictive $t= 0.34$, $p= .631$, and Hybrid $t= 0.86$, $p= .804$, see Appendix G).

A repeated measures ANOVA was conducted to assess whether keyboard layout, text type or other factors were associated with CWPM. The final model for this design included keyboard, age, text type, an age by keyboard interaction, an age by text type interaction, and a text type by keyboard interaction. The model's results indicated that the mean CWPM was not associated with keyboards ($F(2,824)= 1.568$, $p= .209$). However, an age by keyboard effect was significant $F(2,824)= 5.602$, $p= .004$, with participants less than 40 years old achieving greater speeds, on average, with the Hybrid keyboard than the Predictive ($p = .003$) and the Alphabetical ($p= .005$) keyboards. However, this text entry speed advantage was not evident in typing with the Hybrid keyboard rather than the Predictive ($p = .098$) or the Alphabetical ($p = .737$) keyboards for participants who were 40 or more years old.

Furthermore, an age by text type interaction effect was significant ($F(3,824)= 5.553$, $p< .001$). Participants less than 40 years old tended to type faster across all text types than participants who were 40 or more years old. In performing multiple comparisons tests, it was observed that participants less than 40 years old typed Words/Spaces faster, on average, than Sentences ($p < .001$), and typed Sentences faster than Addresses ($p < .001$) and Web text ($p < .001$). Similarly, participants 40 or more years old typed Words/Spaces faster, on average, than Sentences ($p < .001$), and typed Sentences faster than Addresses ($p < .001$) and Web text ($p < .001$).

Additionally, a text type by keyboard interaction effect was significant $F(6, 824)=15.486$, $p < .001$). Specifically, participants typed the Words/Spaces text type, on average, slower using the Alphabetical keyboard than with the Hybrid ($p < .001$) and Predictive ($p < .001$) keyboards. Similarly, participants also typed Sentences, on average, slower using the Alphabetical keyboard than with the Hybrid ($p <.001$) and Predictive ($p =.002$) keyboards. However, participants achieved greater mean text entry speeds using the Alphabetical keyboard than with the Predictive (p <.001) and the Hybrid (p <.001) keyboards when typing Web addresses. Furthermore, participants typed Addresses equally slowly with the Alphabetical keyboard as with the Predictive (p =.624) and the Hybrid (p =.018). It is reasonable to see that the Predictive and Hybrid keyboards did not surpass the Alphabetical keyboard when typing physical Addresses and Web addresses, given that these text types do not conform to the lexical modeling that inherently provides an advantage to the dynamic layouts. The final model results for each case, including the pairwise comparisons, main effects plots, and interaction plots appear in Table 6-16 through 6-19 and Figures 6-12 and 6-13.

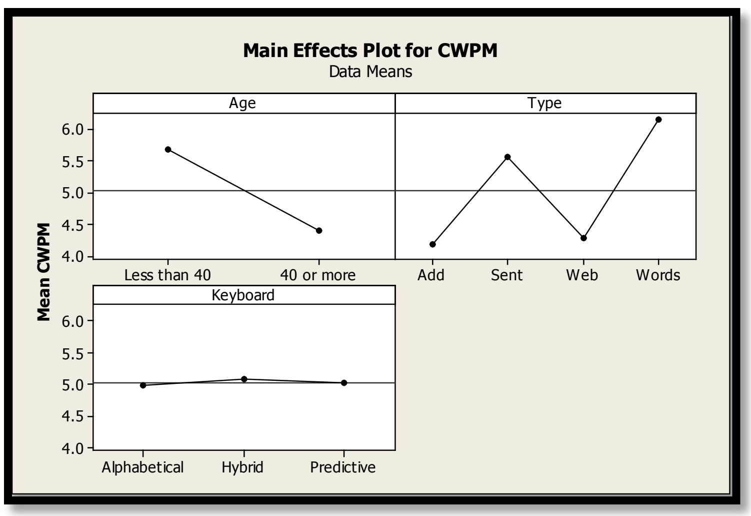


Figure 6-12 Main Effects Plots for CWPM by Age, Text Type, and Keyboard

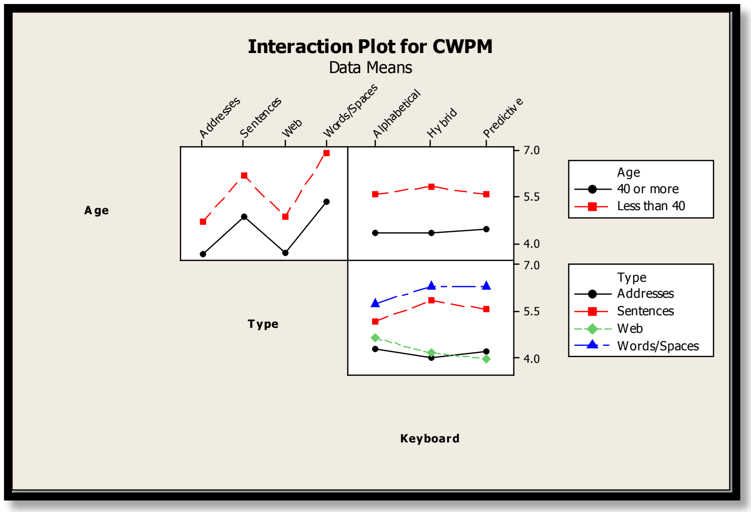


Figure 6-13 Interaction Plots for CWPM by Age, Text Type, and Keyboard

Source	*Numerator df*	*Denominator df*	*F*	*Sig.*
Intercept	1	22	960.218	.000
Age	1	22	15.750	.001
Type	3	824	384.127	.000
Keyboard	2	824	1.568	.209
Age * Type	3	824	5.553	.001
Age * Keyboard	2	824	5.602	.004
Type * Keyboard	6	824	15.486	.000

Table 6-16 Tests of Fixed Effects for CWPM

Age	*(I) Keyboard*	*(J) Keyboard*	*Mean Difference (I-J)*	*Std. Error*	*df*	*Sig.*	*95% Confidence Interval for Difference*	
							Lower Bound	*Upper Bound*
Less than 40	Alphabetical	Predictive	.011	.085	824	.893	-.156	.179
		Hybrid	-.243	.085	824	.005	-.410	-.075
	Predictive	Alphabetical	-.011	.085	824	.893	-.179	.156
		Hybrid	-.254	.085	824	.003	-.422	-.086
	Hybrid	Alphabetical	.243	.085	824	.005	.075	.410
		Predictive	.254	.085	824	.003	.086	.422
40 or More	Alphabetical	Predictive	-.113	.085	824	.188	-.280	.055
		Hybrid	.029	.085	824	.737	-.139	.196
	Predictive	Alphabetical	.113	.085	824	.188	-.055	.280
		Hybrid	.141	.085	824	.098	-.026	.309
	Hybrid	Alphabetical	-.029	.085	824	.737	-.196	.139
		Predictive	-.141	.085	824	.098	-.309	.026

Table 6-17 Age by Keyboard Pairwise Comparisons for CWPM

Age	(I) Type	(J) Type	Mean Difference (I-J)	Std. Error	df	Sig.	95% Confidence Interval for Difference	
							Lower Bound	Upper Bound
Less than 40	Add	Sent	-1.513	.099	824	.000	-1.706	-1.319
		Web	-.154	.099	824	.120	-.347	.040
		Words	-2.230	.099	824	.000	-2.424	-2.037
	Sent	Add	1.513	.099	824	.000	1.319	1.706
		Web	1.359	.099	824	.000	1.166	1.553
		Words	-.718	.099	824	.000	-.911	-.524
	Web	Add	.154	.099	824	.120	-.040	.347
		Sent	-1.359	.099	824	.000	-1.553	-1.166
		Words	-2.077	.099	824	.000	-2.270	-1.883
	Words	Add	2.230	.099	824	.000	2.037	2.424
		Sent	.718	.099	824	.000	.524	.911
		Web	2.077	.099	824	.000	1.883	2.270
40 or More	Add	Sent	-1.236	.099	824	.000	-1.430	-1.043
		Web	-.033	.099	824	.735	-.227	.160
		Words	-1.692	.099	824	.000	-1.886	-1.498
	Sent	Add	1.236	.099	824	.000	1.043	1.430
		Web	1.203	.099	824	.000	1.009	1.397
		Words	-.456	.099	824	.000	-.649	-.262
	Web	Add	.033	.099	824	.735	-.160	.227
		Sent	-1.203	.099	824	.000	-1.397	-1.009
		Words	-1.658	.099	824	.000	-1.852	-1.465
	Words	Add	1.692	.099	824	.000	1.498	1.886
		Sent	.456	.099	824	.000	.262	.649
		Web	1.658	.099	824	.000	1.465	1.852

Table 6-18 Age by Text Type Pairwise Comparisons for CWPM

Type	(I) Keyboard	(J) Keyboard	Mean Difference (I-J)	Std. Error	df	Sig.	95% Confidence Interval for Difference Lower Bound	Upper Bound
Add	Alphabetical	Predictive	.059	.121	824	.624	-.178	.296
		Hybrid	.286	.121	824	.018	.048	.523
	Predictive	Alphabetical	-.059	.121	824	.624	-.296	.178
		Hybrid	.226	.121	824	.061	-.011	.464
	Hybrid	Alphabetical	-.286	.121	824	.018	-.523	-.048
		Predictive	-.226	.121	824	.061	-.464	.011
Sent	Alphabetical	Predictive	-.369	.121	824	.002	-.606	-.132
		Hybrid	-.654	.121	824	.000	-.891	-.417
	Predictive	Alphabetical	.369	.121	824	.002	.132	.606
		Hybrid	-.285	.121	824	.019	-.522	-.047
	Hybrid	Alphabetical	.654	.121	824	.000	.417	.891
		Predictive	.285	.121	824	.019	.047	.522
Web	Alphabetical	Predictive	.669	.121	824	.000	.432	.906
		Hybrid	.481	.121	824	.000	.244	.718
	Predictive	Alphabetical	-.669	.121	824	.000	-.906	-.432
		Hybrid	-.188	.121	824	.120	-.425	.049
	Hybrid	Alphabetical	-.481	.121	824	.000	-.718	-.244
		Predictive	.188	.121	824	.120	-.049	.425
Words	Alphabetical	Predictive	-.562	.121	824	.000	-.799	-.325
		Hybrid	-.541	.121	824	.000	-.778	-.304
	Predictive	Alphabetical	.562	.121	824	.000	.325	.799
		Hybrid	.021	.121	824	.862	-.216	.258
	Hybrid	Alphabetical	.541	.121	824	.000	.304	.778
		Predictive	-.021	.121	824	.862	-.258	.216

Table 6-19 Text Type by Keyboard Pairwise Comparisons for CWPM

6.3.2 Error Rates

6.3.2.1 Uncorrected Error Rate

Accuracy was high overall. The mean uncorrected error rates (UER) across keyboards were 1.5% for Alphabetical, 1.2% for Predictive, and 1.1% for the Hybrid keyboard. The mean error rates across text type were 1.6% for Addresses, 1.5% for Sentences, 1% for

Web addresses, and 1.1% for Words/Spaces. Tables 6-20 and 6-21 provide descriptive statistics for the rate of uncorrected errors by keyboard layout and by text type.

Variable	*Keyboard*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
UER	Alphabetical	288	.015	.039	0	0	0	.017	.333
	Predictive	288	.012	.031	0	0	0	.017	.227
	Hybrid	288	.011	.033	0	0	0	.000	.400

Table 6-20 Descriptive Statistics for Uncorrected Error Rate by Keyboard

Variable	*Text Type*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
UER	Address	216	.016	.032	0	0	0	.017	.211
	Sentences	216	.015	.046	0	0	0	0	.400
	Web	216	.010	.026	0	0	0	0	.200
	Words/Spaces	216	.011	.032	0	0	0	0	.244

Table 6-21 Descriptive Statistics for Uncorrected Error Rate by Text Type

Initially a repeated measures ANOVA was conducted to determine whether keyboard layout or text type was associated with the error rate. The final model for this design included keyboard, text type, gender, experience, age, an age by text type interaction, an age by gender interaction, and a gender by experience interaction. The model results indicated that there was no significant association between UER and keyboard layout ($F(2,832)= 1.329$, $p= .265$) or text type ($F(3,832)= 1.859$, $p= .135$). This supports the idea that although text entry errors are inevitable, these errors were not specifically a consequence of any one of the five-key text entry methods employed in this work. Even so, a gender by experience interaction was associated with UER ($F(1,18)= 10.643$, $p= .004$, see Table 6-22), in which female experts made more mistakes than male experts ($p= .002$) and female non-experts ($p= .003$). Additionally, an age by text type interaction indicating that participants less than 40 years old made fewer errors, on average, than those who were 40 or more years old when typing Words/Spaces (p= .045) and Sentences (p= .013), was suggestive but not significant ($F(3,832)= 3.019$, p= .029).

The UER data, however, was significantly non-normal and zero inflated. Consequently, the test assumptions for the model could not be satisfied (see Appendix G). Various transformations were explored, but none were successful. The pairwise comparisons results and interaction plots appear in Tables 6-23 and 6-24 and Figures 6-14 and 6-15.

Source	*Numerator df*	*Denominator df*	*F*	*Sig.*
Intercept	1	18	121.665	.000
Age	1	18	2.769	.113
Gender	1	18	3.863	.065
Experience	1	18	2.796	.112
Type	3	832	1.859	.135
Keyboard	2	832	1.329	.265
Age * Gender	1	18	3.718	.070
Age * Type	3	832	3.019	.029
Gender * Exp.	1	18	10.643	.004

Table 6-22 Tests of Fixed Effects for Uncorrected Error Rate

Experience	*(I) Gender*	*(J) Gender*	*Mean Difference (I-J)*	*Std. Error*	*df*	*Sig.*	*99% Confidence Interval for Difference* *Lower Bound*	*Upper Bound*
Expert	Female	Male	.012	.003	18	.002	.003	.022
	Male	Female	-.012	.003	18	.002	-.022	-.003
Non-Expert	Female	Male	-.003	.003	18	.371	-.013	.007
	Male	Female	.003	.003	18	.371	-.007	.013

Gender	*(I) Experience*	*(J) Experience*	*Mean Difference (I-J)*	*Std. Error*	*df*	*Sig.*	*99% Confidence Interval for Difference* *Lower Bound*	*Upper Bound*
Female	Expert	Non-Expert	.012	.003	18	.003	.002	.021
	Non-Expert	Expert	-.012	.003	18	.003	-.021	-.002
Male	Expert	Non-Expert	-.004	.003	18	.276	-.013	.006
	Non-Expert	Expert	.004	.003	18	.276	-.006	.013

Table 6-23 Gender by Experience Pairwise Comparisons for Uncorrected Error Rate

Type	*(I) Age*	*(J) Age*	*Mean Difference (I-J)*	*Std. Error*	*df*	*Sig.*	*99% Confidence Interval for Difference* *Lower Bound*	*Upper Bound*
Add	<40	≥40	.006	.005	234.637	.227	-.006	.018
	≥40	<40	-.006	.005	234.637	.227	-.018	.006
Sent	<40	≥40	-.012	.005	234.637	.013	-.024	.000
	≥40	<40	.012	.005	234.637	.013	.000	.024
Web	<40	≥40	.000	.005	234.637	.965	-.012	.012
	≥40	<40	.000	.005	234.637	.965	-.012	.012
Words	<40	≥40	-.009	.005	234.637	.045	-.022	.003
	≥40	<40	.009	.005	234.637	.045	-.003	.022

Table 6-24 Text Type by Age Pairwise Comparisons for Uncorrected Error Rate

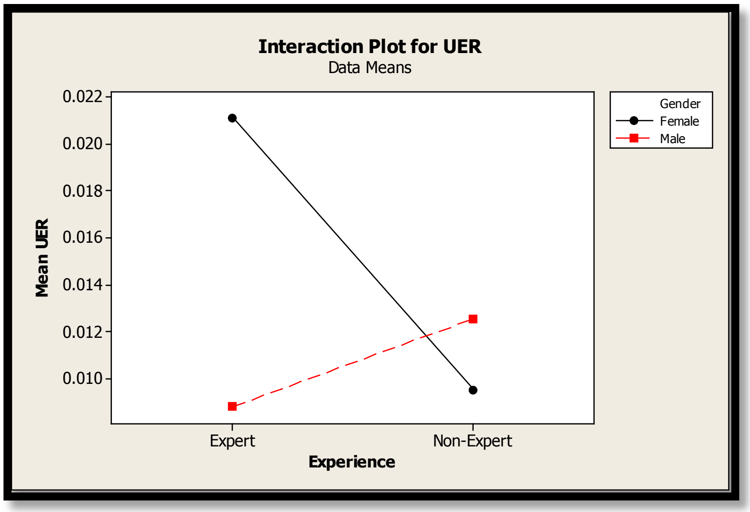


Figure 6-14 Gender by Experience Interaction Plot for Uncorrected Error Rate

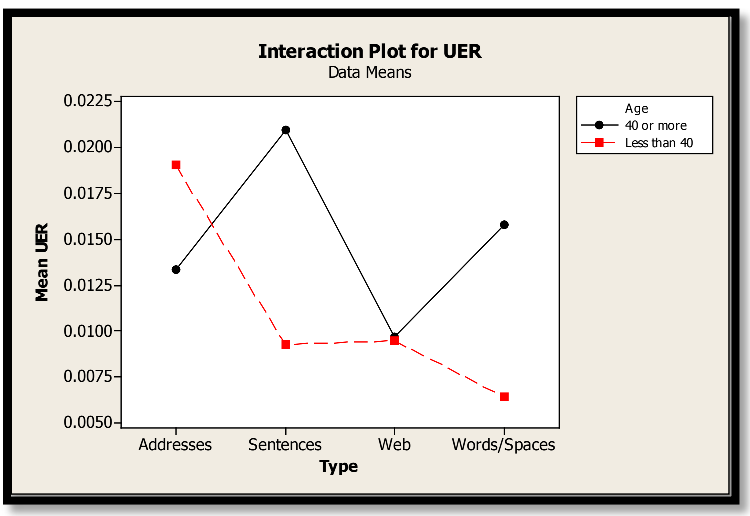


Figure 6-15 Age by Type Interaction Plot for Uncorrected Error Rate

Since the UER data was non-normal and successful transformations could not be applied, Friedman tests were conducted. The test results (Appendix G) indicated that the UER was not associated with keyboard layouts ($X^2(2)$= 0.65, p= .722) or with text types ($X^2(3)$= 6.70, p= .082).

For an alternative analysis, logistic regression incorporating repeated measures was conducted to determine whether keyboard layout or text type was associated with UER. The error rate data was coded to binary outcomes (assigned 0 for no errors occurring within a trial, and 1 for one or more errors occurring within a trial). In total, there were 226 trials in error out of 864.

The final model for the analysis was determined via forward selection procedures, given that using the backward elimination method was not possible due to insufficient degrees of freedom available to run the full factorial model. The final model for this design included keyboard, text type, and a text type by keyboard interaction. Results indicated that the text type by keyboard interaction ($X^2(6) = 27.867$, $p < .001$) was associated with UER, see Table 6-25. Although it is possible that other factors or their interactions may be associated with UER, the data was not sufficiently diversified for further evaluation using Logistic Regression.

Source	*Wald Chi-Square*	*df*	*Sig.*
Intercept	284.110	1	.000
Keyboard	4.394	2	.111
Type	116.509	3	.000
Type*Keyboard	27.867	6	.000

Table 6-25 Tests of Model Effects for Uncorrected Error Rate

6.3.2.2 Total Error Rate

The mean total error rates (TER) across keyboards were 7.6% for Alphabetical, 7.4% for Predictive, and 8.4% for the Hybrid keyboard. The mean TER across text types were 14.3% for Addresses, 8.1% for Sentences, 5% for Web addresses, and 3.8% for Words/Spaces. Tables 6-26 and 6-27 provide descriptive statistics for TER by Keyboard and by Text Type.

Variable	Keyboard	N	M	SD	Min	Q1	Mdn	Q3	Max
TER	Alphabetical	288	.076	.079	0	0	.062	.118	.495
	Predictive	288	.074	.076	0	0	.068	.119	.453
	Hybrid	288	.084	.091	0	0	.062	.137	.464

Table 6-26 Descriptive Statistics for Total Error Rate by Keyboard

Variable	Text Type	N	M	SD	Min	Q1	Mdn	Q3	Max
TER	Address	216	.143	.060	.073	.098	.127	.171	.495
	Sentences	216	.081	.071	0	0	.062	.109	.464
	Web	216	.050	.086	0	0	0	.080	.400
	Words/Spaces	216	.038	.068	0	0	0	.067	.453

Table 6-27 Descriptive Statistics for Total Error Rate by Text Type

A repeated measures ANOVA was conducted to determine whether keyboard layout or text type was associated with TER. The final model for this design, which was identified via forward selection methods, included age, keyboard, and text type. The model results, presented in Table 6-28, indicated that mean TER was associated with text type ($F(3,835)= 99.407$, $p < .001$). Specifically, the error rate was greatest when typing Addresses, over typing Sentences ($p< .001$), Web ($p< .001$), and Words/Spaces ($p< .001$). Similarly, typing Sentences resulted in a greater error rate than typing Web ($p< .001$) and Words/Spaces ($p< .001$, see Figure 6-16 and Table 6-29). The test assumptions for the model, however, could not be satisfied (see Appendix G). Various transformations were explored, but none were successful.

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	22	315.811	.000
Age	1	22	3.318	.082
Type	3	835	99.407	.000
Keyboard	2	835	1.787	.168

Table 6-28 Tests of Fixed Effects for Total Error Rate

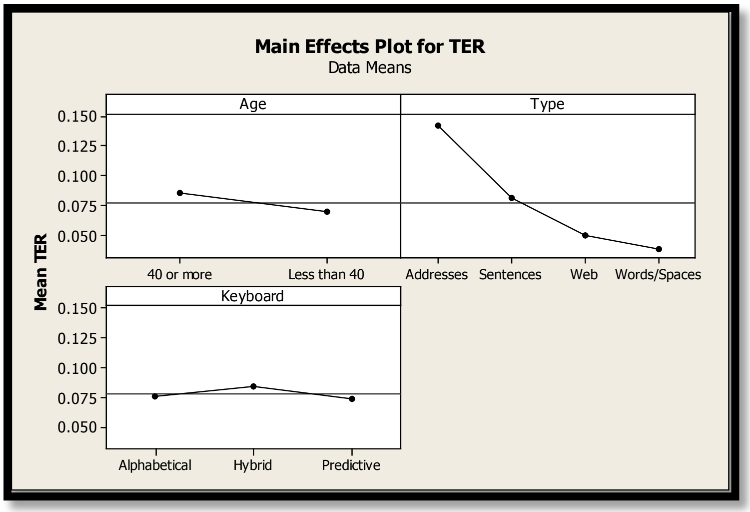


Figure 6-16 Main Effects Plots for Total Error Rate by Age, Text Type, and Keyboard

						99% Confidence Interval for Difference	
(I) Type	*(J) Type*	*Mean Difference (I-J)*	*Std. Error*	*df*	*Sig.*	*Lower Bound*	*Upper Bound*
Add	Sent	.061	.007	835	.000	.040	.082
	Web	.093	.007	835	.000	.072	.114
	Words	.105	.007	835	.000	.084	.126
Sent	Add	-.061	.007	835	.000	-.082	-.040
	Web	.032	.007	835	.000	.011	.053
	Words	.043	.007	835	.000	.022	.064
Web	Add	-.093	.007	835	.000	-.114	-.072
	Sent	-.032	.007	835	.000	-.053	-.011
	Words	.012	.007	835	.477	-.009	.033
Words	Add	-.105	.007	835	.000	-.126	-.084
	Sent	-.043	.007	835	.000	-.064	-.022
	Web	-.012	.007	835	.477	-.033	.009

Table 6-29 Text Type Pairwise Comparisons for Total Error Rate

Friedman tests were conducted given that the TER data was non-normal and successful transformations could not be applied. The test results indicated that the TER was not associated with keyboard layouts ($X^2(2)= 2.33$, $p= .311$). However, TER was associated with text types ($X^2(3)= 61.40$, $p< .001$), see Appendix G. A post-hoc analysis based on Friedman rank-averages indicated that typing Addresses resulted in significantly greater error rates over Sentences ($p < .01$), Web ($p < .0001$), and Words/Spaces ($p < .0001$). Furthermore, typing Sentences resulted in significantly greater errors rates over Words/Spaces ($p < .001$). The remaining rank averages were not significantly different. These findings were consistent with the results from the repeated measures ANOVA.

Logistic regression incorporating repeated measures was also conducted. The error rate data was coded to binary outcomes (assigned 0 for no errors occurring within a trial, and 1 for one or more errors occurring within a trial). In total, there were 579 trials in error out of 864.

The final model for the analysis was determined via forward selection procedures, given that using the backward elimination method was not possible due to insufficient degrees of freedom available to run the full factorial model. The final model for this design included only keyboard layout. Results indicate that keyboard ($X^2(2) = 0.908$, $p = .635$) was not associated with TER (see Table 6-30). Although it is possible that other factors or their interactions may be associated with TER, the data was not sufficiently diversified for further evaluation using Logistic Regression.

Source	*Wald Chi-Square*	*df*	*Sig.*
Intercept	97.142	1	.000
Keyboard	.908	2	.635

Table 6-30 Tests of Model Effects for Total Error Rate

6.3.3 Efficiency

6.3.3.1 KSPC

The overall results for mean KSPC by keyboard were: Alphabetical M= 9.60 (SD= 1.521), Predictive M= 8.0 (SD= 1.998), and Hybrid M= 7.1 (SD= 1.799). Descriptive statistics for KSPC for each by keyboard layout appear in Table 6-31. The values in Table 6-32 were computed and differ from the observed number of keystrokes per character. Figure 6-17 shows both the computed and observed values. The observed values included typematic keystrokes (i.e. counting virtual key-presses during auto-repeat).

Variable	*Keyboard*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
KSPC	Alphabetical	288	9.596	1.521	6.045	8.543	9.432	10.393	15.333
	Predictive	288	7.999	1.998	4.083	6.370	7.767	9.276	14.647
	Hybrid	288	7.104	1.799	3.444	5.423	7.408	8.650	11.063

Table 6-31 Descriptive Statistics for KSPC by Keyboard

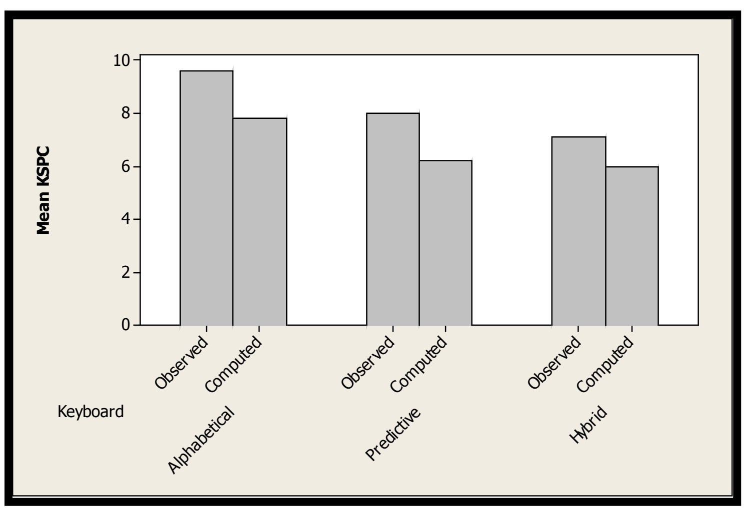


Figure 6-17 Computed and Observed KSPC by Keyboard

Phrase	Type	Computed Average KSPC Alphabetical	Predictive	Hybrid
Mercifully, it was still open.	Sentence	8.40	6.00	4.80
www.flickr.com/explore/	Web	9.13	6.61	7.13
www.digg.com/about/	Web	8.11	7.32	7.37
6129 Lees Pike, 317, Falls Church, VA 22041, jadams@aol.com	Address	7.49	6.66	7.15
www.yelp.com/miami	Web	9.17	8.00	7.94
36 Amber Dr, Pittsford, NY 14534, ravi.adapathya@kodak.com	Address	7.52	7.66	7.45
It ran until past one o'clock.	Sentence	8.47	6.37	5.43
the laser printer is jammed	Words	7.70	4.67	3.33
do not feel too bad about it	Words	6.64	5.25	4.21
I didn't understand why, Clay.	Sentence	7.37	6.13	6.33
where can my little dog be	Words	6.58	5.04	4.19
311 Wembley Rd, Reisterstown, MD 21136, yxaio@umaryland.edu	Address	7.34	6.36	6.95
seasoned golfers love the game	Words	7.50	5.27	3.97
1207 Palo Verde Rd, Irvine, CA 91617, mail@kowym.com	Address	7.88	6.87	7.33
17 Aviation Dr, Winter Haven, FL 33881, jdkochan@aol.com	Address	8.09	6.38	6.64
2 Talbot Pl, Huntington Station, NY 11746, rgulota@tufts.edu	Address	8.28	6.65	6.93
How'd you hear about this one?	Sentence	7.37	5.10	4.87
www.yahoo.com/finance	Web	8.48	6.81	7.71
miami.craigslist.org/mdc/	Web	9.24	8.48	6.40
a big scratch on the tabletop	Words	6.52	5.55	4.55
Oh, that's all right, he said.	Sentence	7.83	6.03	5.10
I'll be waiting for you there.	Sentence	7.63	4.23	4.00
www.travelocity.com/vaca23	Web	7.27	7.35	8.23
never mix religion and politics	Words	8.32	6.10	4.61
www.espn.com/nfl	Web	9.75	8.19	7.81
nothing finer than discovering a treasure	Words	7.71	4.15	3.68
You are all right, my brother?	Sentence	7.53	5.30	5.20
the kids are very excited	Words	5.96	4.32	4.24
I'll just leave the door open.	Sentence	8.27	5.30	4.43
3320 E 68th Ct, Indianapolis, IN 46220, bill@wrbaynes.com	Address	7.79	6.53	7.02
www.giraffe837.com	Web	8.28	7.89	6.78
5303 Foxridge Dr, 301, Mission, KS 66202, daniel@gmail.com	Address	8.05	6.84	7.05
yes you are very smart	Words	5.95	4.36	4.91
5825 Tree Line Dr, Madison, WI 53711, gv@trace.wisc.edu	Address	7.62	6.69	7.16
No, Cady, he made second team.	Sentence	7.13	5.97	5.70
www.wikipedia.org/wiki/Asia	Web	8.30	7.67	7.89

Table 6-32 Computed Minimal KSPC by Keyboard

For the three keyboards, the observed KSPC (including typematic key-presses) was higher than the computed KSPC. Linguistic differences between the specific set of phrases entered and the language model could account for such differences. The effect of this difference depends on the statistical structure of each phrase. This effect, however, should be minor due to the high correlation between the letter frequencies in the phrase set and those in the reference corpus (see Appendix C). The more likely cause of such differences is suboptimal entry, in which case these differences show inefficiency in keyboard usage. As depicted in Figure 6-17, participants entered more keystrokes per character than necessary: 23.2% more for Alphabetical, 27.8% more for Predictive, and 19.4% more for Hybrid. The percent difference was greatest for the Predictive keyboard. This implies that participants did not realize the intended benefits of language modeling with the Predictive keyboard to the same extent as the Hybrid. The most likely reason is that participants tended to overshoot and adjust more often with Predictive than with Hybrid, so participants did more work than necessary with the Predictive keyboard.

The overall results by text type for mean KSPC were: Address M= 8.833 (SD= 1.056), Sentences M= 7.516 (SD= 1.922), Web M= 9.888 (SD= 1.529), and Words/Spaces M= 6.695 (SD= 1.963), as shown in Table 6-33.

Variable	*Text Type*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
KSPC	Address	216	8.833	1.056	6.945	8.166	8.729	9.384	13.411
	Sentences	216	7.516	1.922	4.467	6.067	7.136	8.767	14.385
	Web	216	9.888	1.529	7.033	8.750	9.649	10.778	15.333
	Words	216	6.695	1.963	3.444	5.170	6.149	8.059	13.118

Table 6-33 Descriptive Statistics for KSPC by Text Type

To determine whether keyboard layout, age, texting experience, or gender were associated with average KSPC, a repeated measures ANOVA was conducted. The final model for the analysis was determined via forward selection procedures, since using the backward elimination method was not feasible given insufficient degrees of freedom available to run the full factorial model. The KSPC data was non-normal, but amenable to transformation by the natural log of KSPC plus 1. The final model for this design included the following factors: keyboard, gender, text type and all associated interactions. Model results, as detailed in Table 6-34, indicated a significant text type by keyboard interaction ($F(6,818)= 54.557$, $p<.001$, see Figure 6-18). Post-hoc comparisons (Table 6-35) revealed the following:

- The Alphabetical keyboard requires significantly more keystrokes per character than Hybrid ($p<.001$) and Predictive ($p<.001$) when entering Addresses
- Sentences and Web phrases were entered most efficiently with the Hybrid keyboard over Predictive ($p<.001$) and Alphabetical ($p<.001$)
- The Hybrid keyboard required significantly fewer keystrokes per character than Predictive ($p<.001$) and Alphabetical ($p<.001$) when entering Words/Spaces

These results indicated achievement of the primary design goal of reduction in KSPC for the Hybrid. The KSPC metric was used in this regard to validate design objectives.

Source	*Numerator df*	*Denominator df*	*F*	*Sig.*
Intercept	1	22	34636.780	.000
Type	3	818	442.897	.000
Keyboard	2	818	447.730	.000
Gender	1	22	2.990	.098
Gender * Type	3	818	1.241	.294
Gender * Keyboard	2	818	2.391	.092
Type * Keyboard	6	818	54.557	.000
Gender * Type * Keyboard	6	818	1.900	.078
Dependent Variable: ln(KSPC+1)				

Table 6-34 Tests of Fixed Effects for KSPC

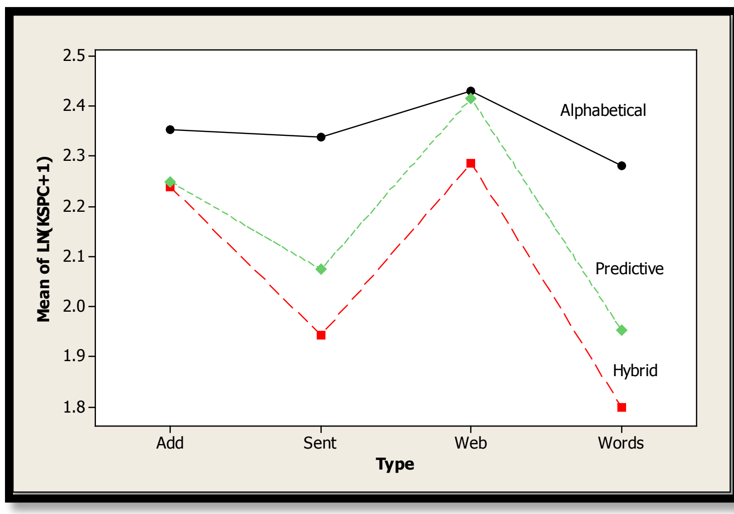


Figure 6-18 Line Plot of Mean KSPC by Keyboard and Text Type

Type	(I) Keyboard	(J) Keyboard	Mean Difference (I-J)	Std. Error	df	Sig.	99% Confidence Interval for Difference Lower Bound	Upper Bound
Add	Alphabetical	Predictive	.115	.022	818	.000	.059	.172
		Hybrid	.126	.022	818	.000	.070	.183
	Predictive	Alphabetical	-.115	.022	818	.000	-.172	-.059
		Hybrid	.011	.022	818	.611	-.045	.067
	Hybrid	Alphabetical	-.126	.022	818	.000	-.183	-.070
		Predictive	-.011	.022	818	.611	-.067	.045
Sent	Alphabetical	Predictive	.297	.022	818	.000	.241	.354
		Hybrid	.450	.022	818	.000	.394	.506
	Predictive	Alphabetical	-.297	.022	818	.000	-.354	-.241
		Hybrid	.153	.022	818	.000	.096	.209
	Hybrid	Alphabetical	-.450	.022	818	.000	-.506	-.394
		Predictive	-.153	.022	818	.000	-.209	-.096
Web	Alphabetical	Predictive	.017	.022	818	.446	-.040	.073
		Hybrid	.160	.022	818	.000	.103	.216
	Predictive	Alphabetical	-.017	.022	818	.446	-.073	.040
		Hybrid	.143	.022	818	.000	.087	.199
	Hybrid	Alphabetical	-.160	.022	818	.000	-.216	-.103
		Predictive	-.143	.022	818	.000	-.199	-.087
Words	Alphabetical	Predictive	.373	.022	818	.000	.317	.430
		Hybrid	.555	.022	818	.000	.499	.611
	Predictive	Alphabetical	-.373	.022	818	.000	-.430	-.317
		Hybrid	.182	.022	818	.000	.125	.238
	Hybrid	Alphabetical	-.555	.022	818	.000	-.611	-.499
		Predictive	-.182	.022	818	.000	-.238	-.125

Dependent Variable: ln(KSPC+1)

Table 6-35 Text Type by Keyboard Pairwise Comparisons for KSPC

6.3.3.2 Movement Inefficiency

Selector movement was also analyzed for movement inefficiency (MI). This measure is indicative of the percentage of keystrokes exceeding the optimal, and is also telling of participants' tendencies to overshoot an intended letter. For this experiment, participants, on average, moved the selector 22.5% more than optimal with the Alphabetical keyboard, 27.1% more with the Predictive, and 19.6% more with the Hybrid. Tables 6-36 and 6-37 provide descriptive statistics for movement inefficiency by keyboard and by text type.

Variable	*Keyboard*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
MI	Alphabetical	288	0.2248	0.163	0	0.111	0.188	0.303	0.926
	Predictive	288	0.2710	0.181	0	0.134	0.245	0.366	0.937
	Hybrid	288	0.1959	0.128	0	0.111	0.177	0.252	1.124

Table 6-36 Descriptive Statistics for Movement Inefficiency by Keyboard

Variable	*Text Type*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
MI	Address	216	0.2162	0.124	0.000	0.129	0.196	0.288	0.691
	Sentences	216	0.2174	0.167	0.000	0.098	0.179	0.302	1.124
	Web	216	0.2566	0.175	0.026	0.124	0.215	0.354	0.937
	Words	216	0.2321	0.173	0.000	0.106	0.203	0.312	0.926

Table 6-37 Descriptive Statistics for Movement Inefficiency by Text Type

A repeated measures ANOVA was conducted. The final model for the analysis was determined via forward selection procedures. It was found that the MI data was non-normal, but amenable to transformation by taking the square root of MI. The final model for this design included keyboard, text type, and a text type by keyboard interaction. The model's results indicated a significant text type by keyboard interaction effect ($F(6,829)= 6.328$, $p< .001$, see Table 6-38 and Figure 6-19). Post-hoc comparisons, shown in Table 6-39, indicated the following:

- Participants were less efficient in selector movement when entering Addresses using the Predictive keyboard than the Hybrid keyboard ($p= .010$)
- Participants were less efficient in selector movement when entering Sentences using the Predictive keyboard than the Hybrid keyboard ($p= .009$)
- Web addresses were entered less efficiently with the Predictive keyboard over Hybrid ($p< .001$) and Alphabetical ($p< .001$)

As is evident, participants tended to overshoot the target character more with the Predictive keyboard than with the Hybrid keyboard.

Source	*Numerator df*	*Denominator df*	*F*	*Sig.*
Intercept	1	23	778.173	.000
Type	3	829	4.338	.005
Keyboard	2	829	16.390	.000
Type * Keyboard	6	829	6.328	.000

Table 6-38 Tests of Fixed Effects for Movement Inefficiency

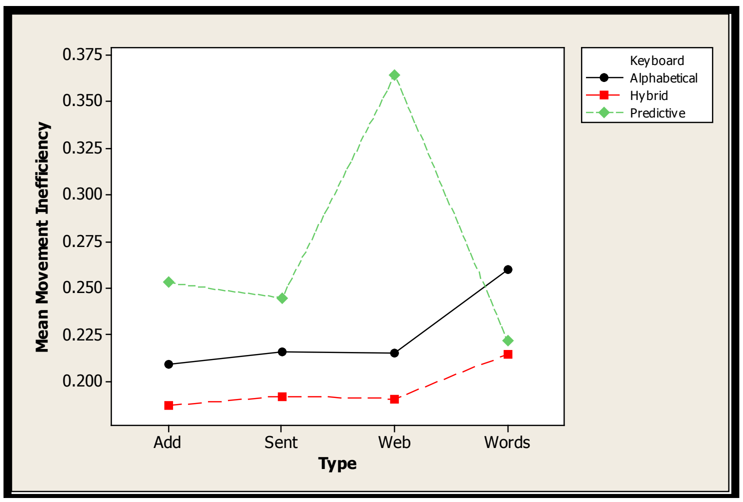


Figure 6-19 Text Type by Keyboard Interaction Plot for Movement Inefficiency

Type	(I) Keyboard	(J) Keyboard	Mean Difference (I-J)	Std. Error	df	Sig.	99% Confidence Interval for Difference	
							Lower Bound	Upper Bound
Add	Alphabetical	Predictive	-.052	.024	829	.028	-.114	.009
		Hybrid	.009	.024	829	.692	-.052	.071
	Predictive	Alphabetical	.052	.024	829	.028	-.009	.114
		Hybrid	.062	.024	829	.010	.000	.123
	Hybrid	Alphabetical	-.009	.024	829	.692	-.071	.052
		Predictive	-.062	.024	829	.010	-.123	.000
Sent	Alphabetical	Predictive	-.029	.024	829	.221	-.091	.032
		Hybrid	.033	.024	829	.160	-.028	.095
	Predictive	Alphabetical	.029	.024	829	.221	-.032	.091
		Hybrid	.063	.024	829	.009	.001	.124
	Hybrid	Alphabetical	-.033	.024	829	.160	-.095	.028
		Predictive	-.063	.024	829	.009	-.124	-.001
Web	Alphabetical	Predictive	-.141	.024	829	.000	-.203	-.080
		Hybrid	.013	.024	829	.575	-.048	.075
	Predictive	Alphabetical	.141	.024	829	.000	.080	.203
		Hybrid	.155	.024	829	.000	.093	.216
	Hybrid	Alphabetical	-.013	.024	829	.575	-.075	.048
		Predictive	-.155	.024	829	.000	-.216	-.093
Words	Alphabetical	Predictive	.046	.024	829	.052	-.015	.108
		Hybrid	.035	.024	829	.138	-.026	.097
	Predictive	Alphabetical	-.046	.024	829	.052	-.108	.015
		Hybrid	-.011	.024	829	.646	-.072	.051
	Hybrid	Alphabetical	-.035	.024	829	.138	-.097	.026
		Predictive	.011	.024	829	.646	-.051	.072

Table 6-39 Text Type by Keyboard Pairwise Comparisons for Movement Inefficiency

6.3.3.3 Typematic Keying

Typematic keying can have a significant impact on text entry (see Tables 6-40 and 6-41). Specifically, for the Alphabetical, Predictive and Hybrid keyboards, typematic events (TE) comprised 58.9%, 47.9%, 35.1% of the observed keystrokes, respectively. For text type, TE comprised 48.5% for Addresses, 46% for Sentences, 53.1% for Web, and 41.6% for Words/Spaces.

Variable	*Keyboard*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
TE	Alphabetical	288	.589	.117	.013	.532	.610	.677	.788
	Predictive	288	.479	.126	.137	.391	.503	.566	.746
	Hybrid	288	.351	.157	.000	.245	.356	.484	.653

Table 6-40 Descriptive Statistics for Typematic Events by Keyboard

Variable	*Text Type*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
TE	Address	216	.485	.123	.077	.406	.510	.582	.682
	Sentences	216	.460	.169	.000	.350	.476	.583	.782
	Web	216	.531	.134	.017	.452	.549	.636	.746
	Words	216	.416	.203	.000	.250	.412	.582	.788

Table 6-41 Descriptive Statistics for Typematic Events by Text Type

There was also a significant relationship between the rates of use of typematic events across the keyboards (see Table 6-42 and Figure 6-20). These relationships indicated that if a participant employed typematic keying with the any one keyboard, then it is likely that participant also employed typematic keying, with similar rates of use, with the other keyboards.

		Alphabetical	Predictive	Hybrid
Alphabetical	Pearson Correlation	1	.440**	.335**
	Sig. (2-tailed)		.000	.000
	N	288	288	288
Predictive	Pearson Correlation	.440**	1	.424**
	Sig. (2-tailed)	.000		.000
	N	288	288	288
Hybrid	Pearson Correlation	.335**	.424**	1
	Sig. (2-tailed)	.000	.000	
	N	288	288	288

***. Correlation is significant at the 0.01 level (2-tailed).*

Table 6-42 Correlation between Typematic Event and Keyboards

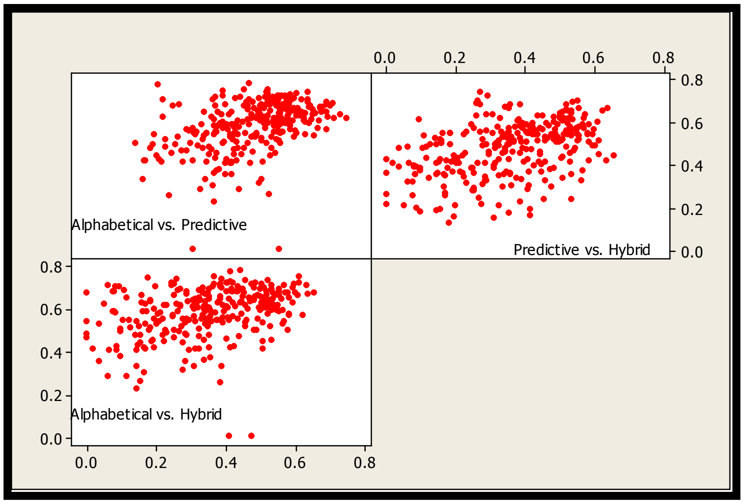


Figure 6-20 Scatter Plots of Keyboard Typematic Events

A repeated measures ANOVA was conducted to determine whether keyboard layout or text type was associated with the rate of use of typematic keying. The final model for the analysis was determined via forward selection procedures. The TE data was non-normal, but amenable to transformation by taking the inverse of the natural log of TE plus 1. The final model for this design included keyboard, text type, and a text type by keyboard interaction. The models results indicated a significant text type by keyboard interaction effect ($F(6,825)= 22.303$, $p<.001$, see Table 6-43 and Figure 6-21). Post-hoc comparison results appear in Table 6-44.

Source	*Numerator df*	*Denominator df*	*F*	*Sig.*
Intercept	1	22.999	397.335	.000
Type	3	825.009	37.035	.000
Keyboard	2	825.018	533.427	.000
Type * Keyboard	6	825.009	22.303	.000

Table 6-43 Tests of Fixed Effects for Typematic Events

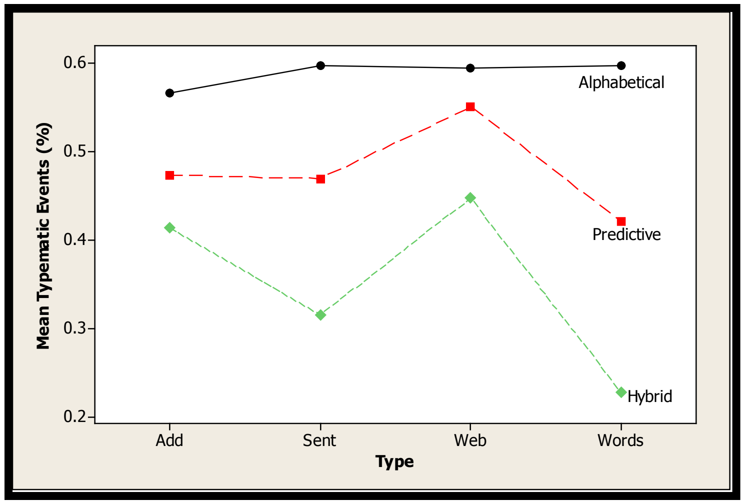


Figure 6-21 Text Type by Keyboard Interaction Plot for Typematic Events

Keyboard	(I) Type	(J) Type	Mean Difference (I-J)	Std. Error	df	Sig.	99% Confidence Interval for Difference Lower Bound	Upper Bound
Alphabetical	Add	Sent	.298	.062	825.001	.000	.137	.459
		Web	.252	.062	825.001	.000	.090	.413
		Words	.344	.062	825.001	.000	.183	.505
	Sent	Add	-.298	.062	825.001	.000	-.459	-.137
		Web	-.046	.062	825.001	.459	-.208	.115
		Words	.046	.062	825.001	.460	-.115	.207
	Web	Add	-.252	.062	825.001	.000	-.413	-.090
		Sent	.046	.062	825.001	.459	-.115	.208
		Words	.092	.062	825.001	.139	-.069	.254
	Words	Add	-.344	.062	825.001	.000	-.505	-.183
		Sent	-.046	.062	825.001	.460	-.207	.115
		Web	-.092	.062	825.001	.139	-.254	.069
Predictive	Add	Sent	-.007	.062	825.001	.916	-.168	.155
		Web	.401	.062	825.001	.000	.240	.563
		Words	-.157	.062	825.001	.012	-.318	.004
	Sent	Add	.007	.062	825.001	.916	-.155	.168
		Web	.408	.062	825.001	.000	.247	.569
		Words	-.150	.062	825.001	.016	-.312	.011
	Web	Add	-.401	.062	825.001	.000	-.563	-.240
		Sent	-.408	.062	825.001	.000	-.569	-.247
		Words	-.558	.062	825.001	.000	-.720	-.397
	Words	Add	.157	.062	825.001	.012	-.004	.318
		Sent	.150	.062	825.001	.016	-.011	.312
		Web	.558	.062	825.001	.000	.397	.720
Hybrid	Add	Sent	-.287	.063	825.012	.000	-.449	-.125
		Web	.135	.062	825.001	.031	-.027	.296
		Words	-.491	.063	825.052	.000	-.654	-.328
	Sent	Add	.287	.063	825.012	.000	.125	.449
		Web	.421	.063	825.012	.000	.260	.583
		Words	-.204	.063	825.021	.001	-.368	-.041
	Web	Add	-.135	.062	825.001	.031	-.296	.027
		Sent	-.421	.063	825.012	.000	-.583	-.260
		Words	-.626	.063	825.052	.000	-.789	-.463
	Words	Add	.491	.063	825.052	.000	.328	.654
		Sent	.204	.063	825.021	.001	.041	.368
		Web	.626	.063	825.052	.000	.463	.789

Dependent Variable: InvLN(TE+1)

Table 6-44 Keyboard by Text Type Pairwise Comparisons for Typematic Events

It was expected that typematic keying would be used more when typing with the Alphabetical keyboard, given its larger KSPC and greater predictability for letter locations. Furthermore, it was also expected that this would hold true across all text types. However, post-hoc tests indicated that typematic keying was used less often when typing Addresses than Sentences ($p<.001$), Web addresses ($p<.001$) and Words/Spaces ($p<.001$) with the Alphabetical keyboard.

Moreover, it was expected to see less typematic keying for the Predictive and Hybrid keyboards as a consequence of their inherently lower KSPCs. Specifically, this should have been true for the Words/Spaces and Sentences text type, which potentially benefits from the lexical modeling employed. However, when typing with the Hybrid keyboard, typematic keying was used less often for typing Words/Spaces than Sentences ($p<.001$), Addresses ($p<.001$) and Web ($p<.001$). A similar effect was observed for the Predictive keyboard, with TE used less when typing Words/Spaces than Web ($p<.001$). Generally, a lower KSPC for the predictive-based layouts should reduce the opportunity for typematic keying. However, with learning, it is possible that the rate of use of TE might increase across all keyboards.

6.3.4 Preference

6.3.4.1 Keyboard Ranking

The mean ranks for keyboard layout were 2.2 for the Alphabetical keyboard, 2.5 for the Predictive keyboard, and 1.4 for the Hybrid keyboard (a lower mean rank is better- closer to first place). The Hybrid keyboard received the most first-place votes (16/24). Table 6-45 and Figure 6-22 show the mean ranking results. Analysis with a Friedman test showed a significant effect of keyboard ($X^2(2)= 15.08$, $p <.001$). A post-hoc analysis based on Friedman rank-averages indicated that participants significantly preferred the Hybrid keyboard over the Predictive keyboard ($p < .001$) and the Alphabetical keyboard ($p < .025$). The remaining rank averages were not significantly different. Table 6-46 contains participants' comments describing the reasons for their rankings.

Keyboard	*Mean Rank*	*Post-Hoc Test*
Alphabetical	2.2	A
Predictive	2.5	A
Hybrid	1.4	B

Notes: Means with same letter are not significantly different according to post-hoc test, with α= .05.

Table 6-45 Mean Keyboard Rankings

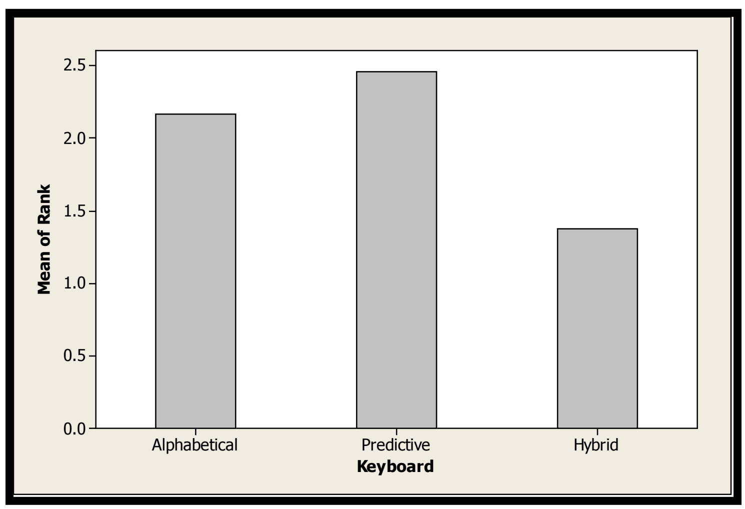


Figure 6-22 Keyboard Layout Preference Ranks

Participant	*Alphabetical*	*Predictive*	*Hybrid*
1	Easy to find letters, but too much scrolling, least favorite because it is too slow	Only good when the target is close to the start position, harder to recognize that using wraparound will help save keystrokes, tiring on the eyes keep glancing back and forth to find letters	Hardest thing was getting to the Caps key, repeated caps are tough, like predictive component and that it combines qualities of the other two keyboards, using predictive component saves some time
2	Hate this keyboard… too much travel required		
3	Easy to find letters… you know where the letters are, but it is very slow	Takes too long to find letters, problematic when you don't see the letter right away	Easiest keyboard, good because it has both (fixed and predictive)
4		Faster with this keyboard, but not knowing where the letters will be is frustrating	
5	This is torture. My wrists are fatigued		
6		Least favorite. "Keyboard anticipates me, but I can't anticipate it"	Typing addresses are tough because predictive component is not very helpful
7		Very frustrating	
8			Caps key needs to be moved
9	This is the one I don't like. So tedious!		
10	understood how to use this one right away, dependable, I don't have to chase the computer, gives comfort	Very frustrating, requires too much visual attention, chasing the computer	More useful
11	Good because I know where the letters will be	Slowest here, not comfortable, only good when letter you want is close	Hate Caps key location. Likes knowing where letters will be (fixed component) and the reduction in travel offered by the predictive component
13		Does not predict that well	
14	Would be better if starting position was in the middle	Like this better than hybrid because majority of targets will be on left half of keyboard. Predictive methods are not great when having to enter the same letter back to back	Going to the delete key takes too long. Predictive does not allow for me to think about how to go to the next letter when traveling to the current letter
15	Easy to use, but it takes longer to get to the letters	May be able to memorize predictive layouts with time, but predictive ability not as strong	Needs getting use to
17		Hate it. It keeps changing, not very useful, not very predictive, the letter you need is never in the first few keys	Easier to use. You don't have to be all over the place searching for the letter you need
18	You know where everything is, but it takes a long time to get there… to much scrolling	If you are not a great speller, this will be really tough to use	Predictive is not easy to use when typing less common words
19	Thumb location and action is very tiring	This is murder, I don't know where the letters will be	Extra thinking is required to find optimal path, it is easier if all in one row
20			Easier than ABC, but need to travel too far to get too caps key
22	Normal, constant rate for typing throughout, limits how fast you can type. Scrolling through the letters is all there is to it, no learning curve.	Too hard to find letters. May be faster for certain words, but it takes too long when prediction fails (i.e. too far off from start)	Reached greatest performance with this one. Can easily develop strategies for use
24	Easier. Letters stay in the same place		Prefer this one. Predictive keys are close to start position and other letters are in a fixed location so you know where to move

Table 6-46 Participants' Comments Associated with Ranking Layouts

6.3.4.2 Keyboard Rating Data

The rating for each layout was the mean of the six items on the Keyboard Layout Rating Form (see section 5.5.4), given after participants finished typing the text phrases with each keyboard. The questionnaire used 7 point scales, with lower ratings better than higher ratings. The mean rating results are in Table 6-47 and Figure 6-23. A repeated measures ANOVA indicated a significant main effect of keyboard layout ($F(2,46)= 9.074$, $p<.001$, see Table 6-48). Post-hoc comparisons showed that participants rated the Predictive keyboard worse than the Alphabetical ($p<.001$) and the Hybrid ($p=.006$) keyboards. The mean rating for the Alphabetical keyboard was not significantly different than the Hybrid keyboard ($p=.207$), see Table 6-49.

Keyboard	*Mean Rating*	*Post-Hoc Test*
Alphabetical	2.85	A
Predictive	4.03	B
Hybrid	3.21	A

Notes: Means with same letter are not significantly different according to post-hoc test, with α= .01. A lower rating is better.

Table 6-47 Mean Keyboard Ratings

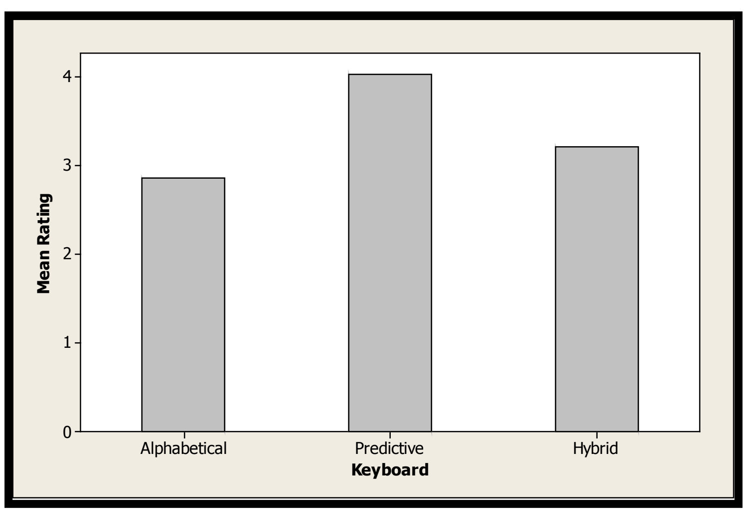


Figure 6-23 Mean Ratings for each Keyboard

Source	*Numerator df*	*Denominator df*	*F*	*Sig.*
Intercept	1	23	403.788	.000
Keyboard	2	46	9.074	.000

Table 6-48 Tests of Fixed Effects for Ratings

(I) Keyboard	*(J) Keyboard*	*Mean Difference (I-J)*	*Std. Error*	*df*	*Sig.*	*99% Confidence Interval for Difference*	
						Lower Bound	*Upper Bound*
Alphabetical	Predictive	-1.173*	.282	46	.000	-1.932	-.415
	Hybrid	-.361	.282	46	.207	-1.119	.397
Predictive	Alphabetical	1.173*	.282	46	.000	.415	1.932
	Hybrid	.812*	.282	46	.006	.054	1.571
Hybrid	Alphabetical	.361	.282	46	.207	-.397	1.119
	Predictive	-.812*	.282	46	.006	-1.571	-.054

Table 6-49 Keyboard Pairwise Comparisons for Rating

There was a lack of correspondence between participant rankings and ratings for the keyboards. The ranking data indicated a clear preference for the Hybrid keyboard over Alphabetical and Predictive. However, this was not fully captured by the keyboard ratings. Pearson correlation coefficients were calculated for all pairs of keyboard rank and ratings (Alphabetical Rank/Rating ($r = .05$), Hybrid Rank/Rating ($r = .30$), and Predictive Rank/Rating ($r = .48$)). Interestingly, the only significant relationship between the rank and rating scores was for the Predictive keyboard ($p = .017$). This discrepancy indicates that one or more aspects of participant preference are not fully captured by the rating questionnaire, suggesting an opportunity for improving it.

6.3.5 Keyboard Attribute Importance

Participants rated all questionnaire attributes (see section 5.5.4) as important, with average importance ratings exceeding 5.5 for all attributes, and with ease of finding letters exceeding 6.0).

6.3.6 Discussion

This experiment addressed the performance at the onset of learning of three alternative keyboard layouts for supporting five-key text entry. Mean text entry throughputs, across text types, were 4.98 CWPM, 5.03 CWPM, and 5.09 CWPM for the Alphabetical, Predictive, and Hybrid keyboards, respectively. The nonstandard keyboards performed better than the Alphabetical keyboard in typing Words/Spaces and Sentences. However, the nonstandard keyboards performed no better for typing standard Address strings. Participants achieved greater mean text entry speeds using the Alphabetical keyboard than with the nonstandard keyboards when typing Web addresses. This is most likely an effect of the optimization strategy, based on English digraphs, employed in designing the predictive layouts. Even so, it appears that performance did not attenuate as much, across text types, for the Alphabetical keyboard as with the Hybrid and Predictive keyboards. All keyboards followed the same basic pattern in performance across text types, but not to the same degree.

Results also suggest that using the Hybrid or Alphabetical keyboards at the novice level will not lead to an overall reduction in error rate in comparison to using the Predictive keyboard. Although accuracy was high overall, typing Addresses resulted in higher total error rates across keyboards. Participants, however, took time to correct the errors they made, as is evident by the differences in the two types of error rates. This finding is reasonable because it is common for corrected errors to vastly outnumber uncorrected errors. Overall, it appears that none of the methods studied led to a significant gain in text entry rate or a reduction in error at this level of user training. Despite a lack of a strong association among the three keyboard layouts on the dependent measures for speed and error rates, there were substantial differences between the three methods on the other dependent measures.

Keyboard designs were optimized to reduce KSPC. In this evaluation, the Hybrid keyboard had the lowest difference in observed vs. computed KSPC than the Predictive and Alphabetical keyboards. Further, the analysis indicated that most text types can be entered with fewer keystrokes, on average, when using the Hybrid keyboard.

Using the Predictive keyboard resulted in the greatest inefficiency in selector movement. Results indicated that participants tended to overshoot the target character more with the Predictive keyboard than with the Hybrid keyboard.

As expected, given the larger KSPC and the greater predictability for letter locations, typematic keying was used at a greater rate with the Alphabetical keyboard. However, opportunities for typematic keying are particularly interesting for these methods overall because cursor distances are in some cases substantial. It is reasonable to expect that, with learning, the rate of use of TE would increase across all keyboards.

When raking layouts, the Hybrid keyboard received the most first-place votes (16/24). Participants' comments indicated that the advantages of the digram-based layout are more accessible when typing with the Hybrid Keyboard. Furthermore, participants' comments indicated that the Predictive keyboard was the most frustrating of the methods due to the high visual attention required.

6.4 Experiment #3: Expert Performance

6.4.1 Missing Trials

In this experiment, a trial was the entry of one test phrase. Overall, there were 2160 trials, as all participants completed 9 trials at each of 20 sessions. Data log files for 11 trials were corrupted. CWPM data for all missing trials were recuperated via analysis of session videos. Unfortunately, the data recovery was not possible for the other performance measures.

6.4.2 Text Entry Speed

The mean entry rates across all twenty sessions were 6.874 (SD= 1.51) for Alphabetical, 7.737 (SD= 1.87) for Predictive, and 7.794 (SD= 1.90) for Hybrid. At the last session, mean entry rates were 7.728 (SD= 1.41) for Alphabetical, 8.801 (SD= 1.79) for Predictive, and 8.929 (SD= 1.89) for Hybrid. Maximum session averages were 10.612 for Alphabetical at Session 13, 13.169 at Session 13 for Predictive, and 14.039 at Session 20 for Hybrid. Tables 6-50 and 6-51 provide descriptive statistics for text entry rates by keyboard layout across all sessions and at Session 20, respectively. Figure 6-24 depicts the mean entry rates by session and keyboard layout.

Variable	*Keyboard*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
$CWPM_{All}$	Alphabetical	720	6.874	1.505	1.492	5.850	6.866	8.025	10.612
	Predictive	720	7.737	1.871	2.949	6.415	7.697	9.074	13.169
	Hybrid	720	7.794	1.902	2.572	6.417	7.798	9.142	14.039

Table 6-50 Descriptive Statistics for Text Entry Rates across all Sessions

Variable	*Keyboard*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
$CWPM_{20}$	Alphabetical	36	7.728	1.411	4.997	6.631	7.798	9.131	10.454
	Predictive	36	8.801	1.792	4.823	7.694	9.005	9.852	12.242
	Hybrid	36	8.929	1.890	4.853	7.890	8.705	10.275	14.039

Table 6-51 Descriptive Statistics for Text Entry Rates at Session 20

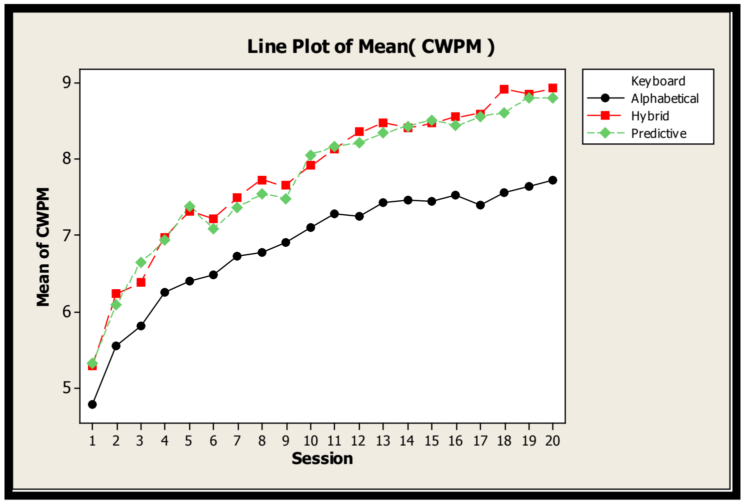


Figure 6-24 Text Entry Rates for the Alphabetical, Predictive, and Hybrid Keyboards across Sessions

As in the previous experiments, there was a significant positive relationship in the text entry rates across the keyboards (see Table 6-52 and Figure 6-25). These results indicated that if a participant performed well with one keyboard, then it was likely that participant also performed well with the other keyboards.

		Alphabetical	Predictive	Hybrid
Alphabetical	Pearson Correlation	1	.826**	.856**
	Sig. (2-tailed)		.000	.000
	N	720	720	720
Predictive	Pearson Correlation	.826**	1	.830**
	Sig. (2-tailed)	.000		.000
	N	720	720	720
Hybrid	Pearson Correlation	.856**	.830**	1
	Sig. (2-tailed)	.000	.000	
	N	720	720	720

**. Correlation is significant at the 0.01 level (2-tailed).

Table 6-52 Correlation between Text Entry Rate and Keyboard

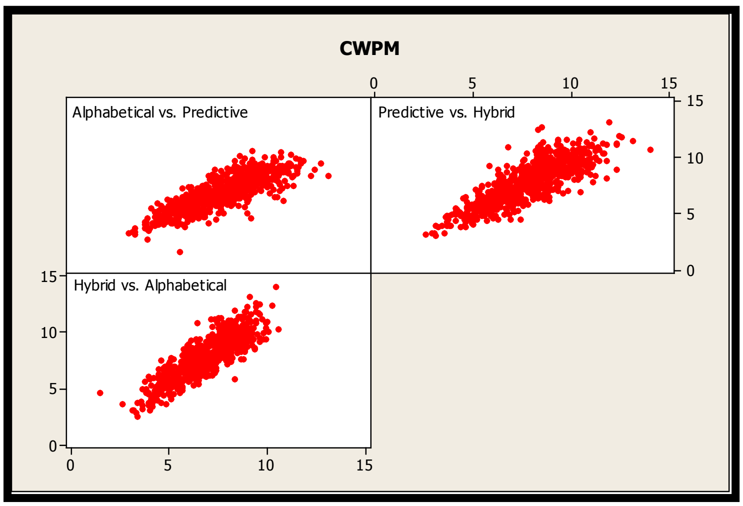


Figure 6-25 Scatter Plots of Keyboard Text Entry Rates (CWPM)

For this experiment, one hypothesis was that expert entry speeds for the Hybrid keyboard would be greater than or equal to 15 CWPM. The descriptive statistics provided above clearly show that the observed average entry speeds for all of the three keyboards was not even greater than 10 CWPM. In fact, the maximum entry speed in the entire study, across all keyboards, was ~14 CWPM for the Hybrid keyboard. Therefore, the evidence did not support this hypothesis.

It was also hypothesized that the entry rates for the Hybrid keyboard would exceed those for the Alphabetical and Predictive keyboards after 90 minutes of practice (~ at Session 10). Analysis with pairwise t-tests for Session 10, however, indicated that there was no significant difference in text entry rates between the Hybrid and Predictive keyboards ($t(35)= 0.620$, $p=.539$). In partial support of the hypothesis, the Alphabetical keyboard had significantly slower text entry rates at Session 10 than the Predictive ($t(35)= -5.165$, $p<.001$) and the Hybrid ($t(35)= 5.138$, $p<.001$) keyboards.

6.4.2.1 Model Fitting

Descriptive exploratory analyses of the data (specifically, empirical growth plots for text entry learning rates for each participant) were conducted prior to fitting statistical models. Figure 6-26 shows the relationship between entry rates and session, across keyboards for each participant. The participants had distinct change trajectories.

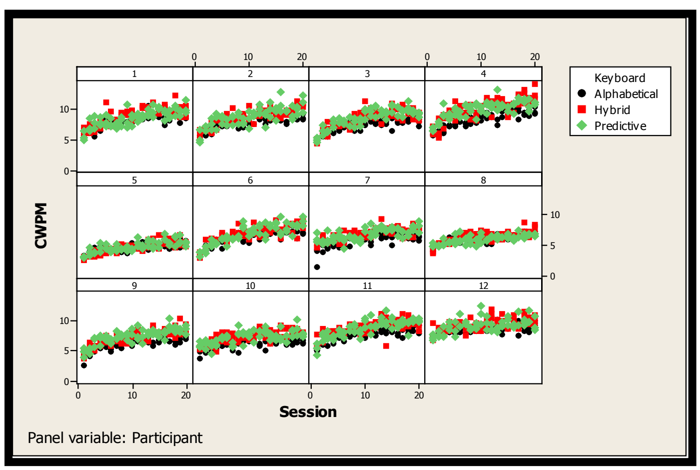


Figure 6-26 Participants' Text Entry Rates across Sessions

The S-Plus statistical software package was used to fit the best model to the data of the individual learning profiles. During the model fitting process, various residual plots were monitored to check for departures from model assumptions; such as non-constant variance, non-linearity in the residuals, non-normality or substantial auto-correlations associated with repeated measurements.

Akaike information criterion (AIC) and Bayesian (Schwarz) information criterion (BIC) goodness-of-fit statistics as generated by S-Plus were used to compare different models during the model fitting process. Smaller values of AIC and BIC criterion indicate a better model fit across models with the same fixed effects. However, these values have to be considered in conjunction with the various residual plots to determine the best model fit.

6.4.2.1.1 Power law

Starting with a power model and using a forward selection approach, aided by residual plots and goodness-of-fit criteria, the following mixed effects model was selected for text entry learning:

(1) $Log(CWPM) = 1.548166 + 0.115059Keyboard_P + 0.120458Keyboard_H + 0.166114*Log(Session)$

Therefore, the estimated average equations for the learning curves for each keyboard were:

(2) $Log(CWPM_{Alphabetical}) = 1.548166 + 0.166114*Log(Session)$

(3) $\text{Log}(\text{CWPM}_{\text{Predictive}}) = 1.663225 + 0.166114*\text{Log}(\text{Session})$

(4) $\text{Log}(\text{CWPM}_{\text{Hybrid}}) = 1.668624 + 0.166114*\text{Log}(\text{Session})$

Correlation within participants was captured via an auto regressive model with lag 1 (AR1). Indicator variables were chosen for keyboard as the learning curves obtained were specific to the levels selected. Keyboard had three levels that were modeled by using two indicator variables, namely Keyboard_P and Keyboard_H for the Predictive and Hybrid keyboards, respectively. The coefficients for these variables were compared to the Alphabetical keyboard, which was the baseline. The model only incorporated the effect of Keyboard and Session on text entry rates (CWPM), as age, gender, and experience did not have a significant effect on learning.

The response versus fitted value plot appears in Figure 6-27. The plot shows a good model fit to the skill acquisition data. The standard deviations for the random effect terms are in Table 6-53. Collectively, standard deviations of the intercept and session random effects, which represent the variability in participant-specific adjustments from the overall average model, explain a large portion of the variability in the observed data when compared to the standard deviation of the error residual term. This indicates that the variation between participants under similar conditions was greater compared to the within participant variation, as captured by the standard deviation of the residual error. Furthermore, there is evidence that inter-participant standard variations differed across keyboards, which implies that variability in performance among participants may be keyboard dependent. Specifically, inter-participant standard deviations were lower for typing with the Alphabetical than the nonstandard keyboards (see Appendix H).

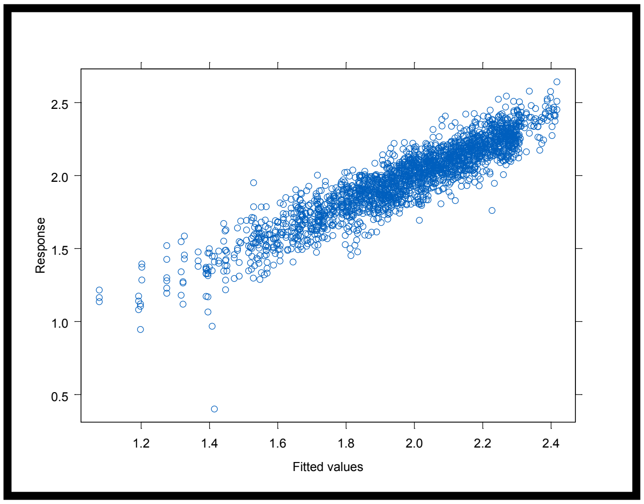


Figure 6-27 Response vs. Fitted Value Plot of the Power Law Model for Text Entry Rate

Parameter	*Std Dev*	*Correlation Matrix*	
		Intercept	**Session**
Intercept	0.237682		
Session	0.040667	-0.601	
Residual	0.108818		

Table 6-53 Estimates of the Variance Components of the Random Effect Terms of the Power Law Model

Table 6-54 shows the model parameter estimates and the associated *p*-values. The average initial entry rates for keyboard layouts studied were 4.70 for the Alphabetical, 5.28 for the Predictive, and 5.30 for the Hybrid. Among the main effects, the coefficients for $Keyboard_P$ (t(2145)= 19.819, p < .0001) and $Keyboard_H$ (t(2145)= 21.331, p < .0001) were positive and highly significant. However, initial entry rates did not differ between the Hybrid and Predictive keyboards (F(1,2145)= 0.782, p = .376).

This implies that average initial entry rates when typing with the nonstandard keyboards tend to be greater than the average initial entry rates when typing the Alphabetical keyboard.

Parameter	*Value*	*Std. Error*	*DF*	*t-value*	*p-value*
Intercept	1.5482	0.0691	2145	22.4166	<.0001
Log(Session)	0.1661	0.0122	2145	13.6397	<.0001
$Keyboard_P$	0.1151	0.0058	2145	19.8190	<.0001
$Keyboard_H$	0.1205	0.0056	2145	21.3308	<.0001

Table 6-54 Parameter Estimates of the Power Model

Figure 6-28 presents the estimated profiles based on the fitted model. Learning rates are not captured in the parameters for this model. As shown in Figure 6-28, the effect of including keyboard in the model was shifting the curve up or down, but it did not change the learning rate. Therefore, all three keyboard curves within each participant profile were parallel to each other and only differed by the intercept. As is evident, the model results did not indicate a crossover point (where one technique after sufficient practice outperforms another) for entry rate.

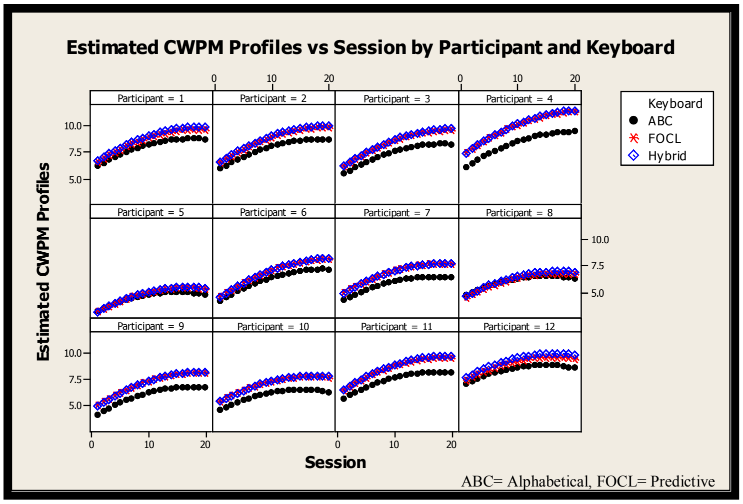


Figure 6-28 Power Law Model Estimated Learning Trends for Text Entry Rates

6.4.2.1.2 Expanded Power Law Model

When modeling with the power law function, as described previously, adding Keyboard as a random effect led to numerical difficulties in variance estimations. In not being able to include Keyboard as a random effect, inter-individual differences in terms of keyboards studied cannot be assessed. Furthermore, learning rates for each keyboard could not be evaluated as a keyboard by session interaction is not a part of the usual power law model. To overcome these limitations, an expanded power law model incorporating this interaction was fitted to the data.

The final form of the expanded power law model was selected using a forward selection procedure and by inspection of associated residual plots and goodness-of-fit criteria. No non-linear models were considered due to their sensitivity to lack of fit. The following mixed effects model was selected for text entry learning:

(1) $Log(CWPM)= 1.565237 + 0.097562*Keyboard_P + 0.083796*Keyboard_H + 0.158178*Log(Session) + 0.008011*Keyboard_P*Log(Session) + 0.017258*Keyboard_H*Log(Session)$

Therefore, the estimated average equations for the learning curves for each keyboard were:

(2) $Log(CWPM_{Alphabetical})= 1.565237 + 0.158178*Log(Session)$

(3) $Log(CWPM_{Predictive})= 1.662799 + 0.166189*Log(Session)$

(4) $Log(CWPM_{Hybrid})= 1.649033 + 0.175436*Log(Session)$

Correlation within participant was captured via an auto regressive model with lag 1 (AR1). Indicator variables were chosen for keyboard as the learning curves obtained are specific to the levels selected. Keyboard had three levels that were modeled by using two indicator variables, namely $Keyboard_P$ and $Keyboard_H$ for the Predictive and Hybrid keyboards, respectively. The coefficients for these variables were compared to the baseline Alphabetical keyboard. As in the power law model, this model only incorporated the effect of Keyboard and Session on text entry rates (CWPM). Age, gender, and experience did not have a significant effect on learning.

The response versus fitted value plot appears in Figure 6-29. The plot shows a good model fit to the skill acquisition data. The standard deviations for the random effect terms are in Table 6-55. The combined standard deviations of the intercept, keyboard, and session random effects explain a large portion of the variability in the observed data when compared to the standard deviation of the error residual term. As a result, the variation between participants under similar conditions is greater compared to the within participant variation, as captured by the standard deviation of the residual error. Table 6-55 also shows that there was a low correlation between session and Predictive keyboard effects (0.252) and Hybrid keyboard effects (0.039), but there was a high correlation (0.967) between the Predictive and Hybrid keyboard effects.

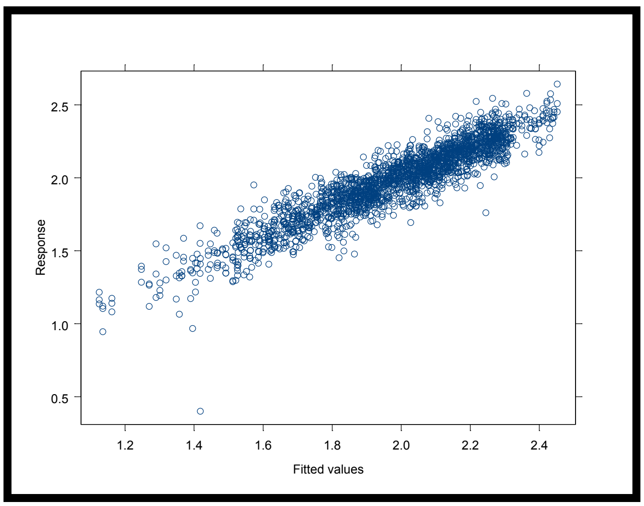


Figure 6-29 Response vs. Fitted Value Plot of the Expanded Power Law Model

Parameter	*Std Dev*	*Correlation Matrix*		
Intercept	0.222310	**Intercept**	**Session**	**Keyboard$_P$**
Session	0.039520	-0.633		
Keyboard$_P$	0.053724	-0.012	0.252	
Keyboard$_H$	0.047768	0.233	0.039	0.967
Residual	0.098434			

Table 6-55 Estimates of the Variance Components of the Random Effect Terms of the Expanded Power Law Model

Table 6-56 shows the model parameter estimates and the associated *p*-values. The average initial entry rates for keyboard layouts studied were 4.78 for the Alphabetical, 5.27 for the Predictive, and 5.20 for the Hybrid. Among the main effects, the coefficients for $Keyboard_P$ (t(2143)= 4.389, p <.0001) and $Keyboard_H$ (t(2143)= 3.978, p< .0001) were positive and highly significant. Initial entry rates did not differ between the Hybrid and Predictive keyboards (F(1,2143)= 0.702, p= 0.402). Therefore, average initial entry rates when typing with the nonstandard keyboards were generally higher than the average initial entry rates when typing the Alphabetical keyboard.

Parameter	*Value*	*Std. Error*	*DF*	*t-value*	*p-value*
Intercept	1.565237	0.065284	2143	23.975	<.0001
Log(Session)	0.158178	0.012505	2143	12.649	<.0001
$Keyboard_P$	0.097562	0.022306	2143	4.387	<.0001
$Keyboard_H$	0.083796	0.021064	2143	3.978	<.0001
$Keyboard_P$*Log (Session)	0.008011	0.007037	2143	1.138	0.2551
$Keyboard_H$*Log(Session)	0.017258	0.007038	2143	2.452	0.0143

Table 6-56 Parameter Estimates of the Expanded Power Model using Session

In the expanded power model, learning rates are captured in the model parameters via the keyboard by session interactions. Trend in time data indicated a marginally significant difference in learning rates between the Hybrid and Alphabetical keyboards (t (2143) =2.452, p=.014). However, trend in time data did not suggest a significant difference in learning rates between Predictive and Alphabetical (t (2143) =1.138, p=0.255) or between Predictive and Hybrid (F(1,2143)=1.727, p=.189). This implies that the average change in entry rates when typing with the Hybrid keyboard was generally higher than the average change in entry rates when typing with the Alphabetical keyboard, but average changes in entry rates did not differ across the nonstandard keyboards.

Figure 6-30 presents the estimated profiles based on the fitted model. Based on the model results, there is no crossover point for entry rate due to inferior entry rates for the Alphabetical keyboard. This finding is supported by the positive parameter estimates for Predictive and Hybrid keyboards and expressed in Figure 6-30.

Overall, text entry rates increased over time for all participants. For some, the increase was more rapid than for others. Furthermore, leveling off of text entry rates in time is not apparent in the figure. This implies that even greater text entry rates for the keyboards could be attained with more time. This experiment does not allow claims for identification of an upper limit given the overall small sample size and that an asymptote was not apparent in the 20 sessions that made up the follow-up time for this experiment.

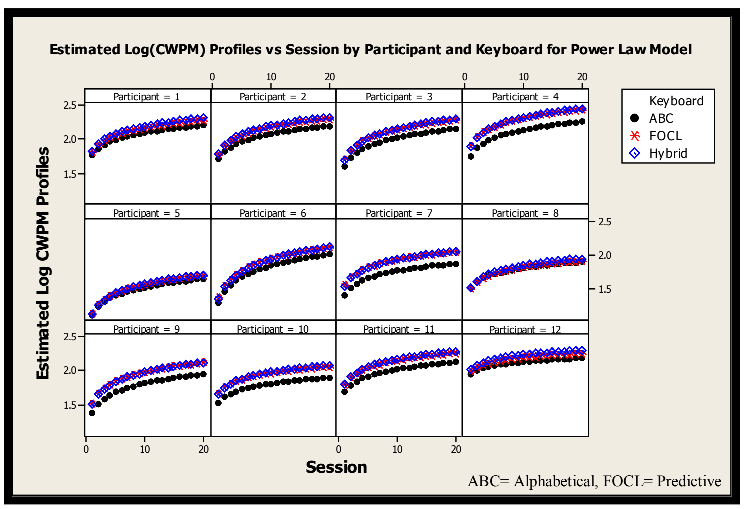


Figure 6-30 Expanded Power Law Model Estimated Learning Trends for Text Entry Rates

6.4.2.1.3 Quadratic Model

The best fitting model to the data, among the models considered, was a quadratic model. The best quadratic model was selected using a forward selection procedure and by inspection of associated residual plots and goodness-of-fit criteria. No non-linear models were considered due to their sensitivity to lack of fit.

The following mixed effects quadratic model was selected for text entry learning:

(1) $CWPM = 4.876023 + 0.565791*Keyboard_P + 0.547498*Keyboard_H + 0.314311*Session + 0.029514*Keyboard_P*Session + 0.036262*Keyboard_H*Session - 0.009125*Session^2$

Therefore, the estimated average equations for the learning curves for each keyboard were:

(2) $CWPM_{Alphabetical} = 4.876023 + 0.314311*Session - 0.009125*Session^2$

(3) $CWPM_{Predictive} = 5.441814 + 0.343825 *Session - 0.009125*Session^2$

(4) $CWPM_{Hybrid} = 5.423521 + 0.350573*Session - 0.009125*Session^2$

Correlation within participant was captured via an auto regressive model with lag 1 (AR1). Indicator variables were chosen for keyboard as the learning curves obtained are specific to the levels selected. Keyboard had three levels, modeled by using two indicator variables, $Keyboard_P$ and $Keyboard_H$ for the Predictive and Hybrid keyboards, respectively. The coefficients for these variables were compared to the baseline Alphabetical keyboard. As in the power law models, the quadratic model only incorporated the effect of keyboard and session on text entry rates (CWPM). Age, gender, and experience did not have a significant effect on learning.

The response versus fitted value plot appears in Figure 6-31. The plot shows a good model fit to the skill acquisition data. The standard deviations for the random effect terms are in Table 6-57. Standard deviations of the intercept random effect term as well as the session and keyboard random effects explained a large portion of the variability in the observed data when compared to the standard deviation of the error residual term. Therefore, the variation between participants under similar conditions was greater compared to the within participant variation, as captured by the standard deviation of the residual error. Furthermore, there is evidence that inter-participant standard variations differed across keyboards, which implies that variability in performance among participants may be keyboard dependent. Specifically, inter-participant standard deviations were lower for typing with the Alphabetical than the nonstandard keyboards (see Appendix H). Table 6-57 also shows that there was a moderate correlation between session and Predictive keyboard effects (0.604) and Hybrid keyboard effects (0.565). There was a high correlation (0.973) between the Predictive and Hybrid keyboard effects.

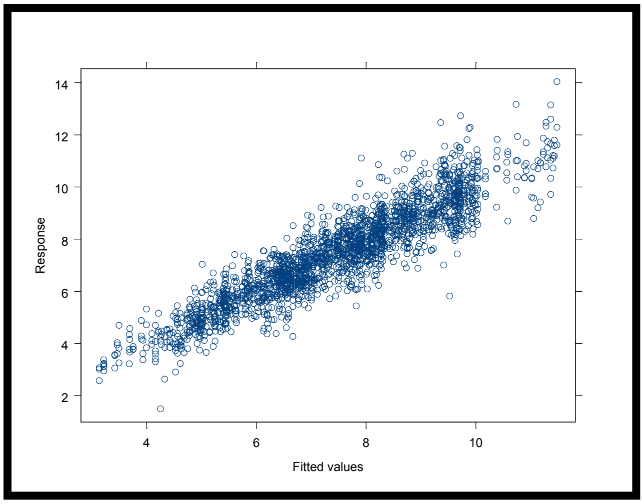


Figure 6-31 Response vs. Fitted Value Plot of the Quadratic Model for Text Entry Rate

Parameter	***Std Dev***	***Correlation Matrix***		
		Intercept	**Session**	**Keyboard$_P$**
Intercept	1.133079			
Session	0.034684	0.078		
Keyboard$_P$	0.437755	0.222	0.604	
Keyboard$_H$	0.403469	0.441	0.565	0.973
Residual	0.825155			

Table 6-57 Estimates of the Variance Components of the Random Effect Terms of the Quadratic Model

Table 6-58 shows the model parameter estimates and the associated *p*-values. The average initial entry rates for keyboard layouts studied were 4.88 for the Alphabetical, 5.44 for the Predictive, and 5.42 for the Hybrid. Among the main effects, the coefficients for Keyboard$_P$ ($t(2142)= 3.749$, $p< .001$) and Keyboard$_H$ ($t(2142)= 3.868$, $p< .001$) were positive and highly significant. However, initial entry rates did not differ between the Hybrid and Predictive keyboards ($F(1,2142)= 0.033$, $p = 0.856$). This means that average initial entry rates when typing with the nonstandard keyboards were generally higher than the average initial entry rates when typing the Alphabetical keyboard.

Parameter	Value	Std. Error	DF	t-value	p-value
Intercept	4.876023	0.333163	2142	14.63554	<.0001
Session	0.314311	0.016030	2142	19.60758	<.0001
$Session^2$	-0.009125	0.000567	2142	-16.07659	<.0001
$Keyboard_P$	0.565791	0.150896	2142	3.74955	<.0001
$Keyboard_H$	0.547498	0.141529	2142	3.86844	<.0001
$Keyboard_P$* Session	0.029514	0.006881	2142	4.28923	<.0001
$Keyboard_H$*Session	0.036262	0.006711	2142	5.40340	<.0001

Table 6-58 Parameter Estimates of the Quadratic Model using Session

Learning rates are captured in the model parameters via the keyboard by session interactions. Trend in time data indicated a significant difference in learning rates between the Predictive and Alphabetical keyboards ($t(2142) = 4.289$, $p < .001$) and between the Hybrid and Alphabetical keyboards ($t(2142) = 5.403$, $p < .001$). However, trend in time data did not suggest a significant difference in learning rates between the Predictive and Hybrid keyboards ($F(1,2142) = .710$, $p = .399$). This means that average change in entry rates when typing with the nonstandard keyboards were generally higher than the average change in entry rates when typing the Alphabetical keyboard, but average changes in entry rates did not differ between the nonstandard keyboards.

Furthermore, based on the model results, there is no crossover point for entry rate, specifically because throughout the study performance with the Alphabetical keyboard was inferior to the Hybrid and Predictive keyboards and because there was no significant difference between the nonstandard keyboards. The positive parameter estimates for the Predictive and Hybrid keyboards support this finding.

Figure 6-32 presents the estimated profiles based on the fitted model. Text entry rates increased over time for all participants. For some, the increase was more rapid, for others, less so. Furthermore, no leveling off of text entry rates in time is apparent in the figure, and the data do not support identification of an upper limit. As in the expanded power model, it may be possible that greater text entry rates for the keyboards could be attained with time.

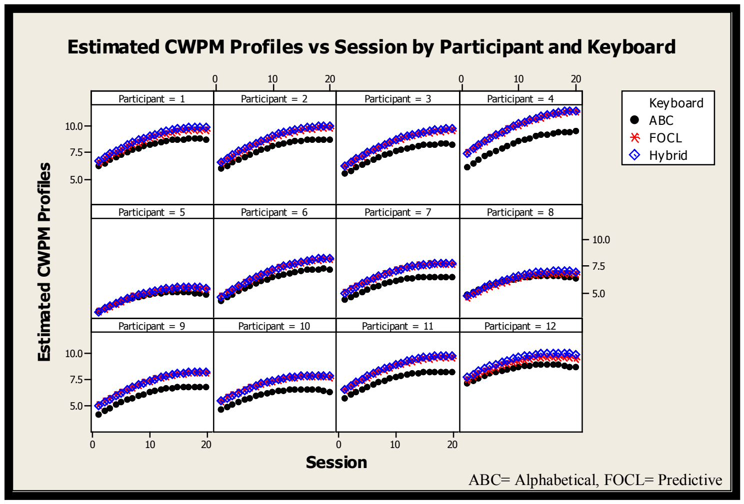


Figure 6-32 Quadratic Model Estimated Learning Trends for Text Entry Rates

6.4.2.1.4 Power vs. Quadratic Models

The power and quadratic models presented above were identified via empirical approaches. The power models were employed mainly due their ubiquity in text entry research. As described previously, adding Keyboard as a random effect in the preliminary power model led to numerical difficulties in variance estimations. Also, the preliminary power law cannot not be used to evaluate learning rates for each keyboard as this model does not incorporate keyboard by session interaction. The data modeled with an expanded power law overcame these restrictions. However, the expanded power model underestimated expert rates in comparison to observed behavior. The quadratic model, on the other hand, did not suffer these stated limitations and had superior residual behavior. For these reasons, the quadratic model seems to provide the best fit for the skill acquisition data herein and is therefore the preferred model for this work.

6.4.2.1.5 Cumulative Time vs. Session

In the models presented above, time was measured in sessions. However, the data collection schedule was flexible in that each individual had a unique schedule. Therefore, test sessions for each participant were not equally spaced, as is suggested when modeling learning rates using session data. Using actual time, measured in hours, might increase the precision of the model. Thus, an analysis using actual time was conducted to determine if the unequally spaced sessions had an effect on model results.

The best model was selected using a forward selection procedure and by inspection of associated residual plots and goodness-of-fit criteria. The following mixed effects quadratic model was selected for text entry learning:

$$(1)\ CWPM = 5.395594 + 0.629198*Keyboard_P + 0.612894*Keyboard_H + 0.226239*Time + 0.025383*Keyboard_P*Time + 0.032037*Keyboard_H*Time - 0.005125*Time^2$$

Therefore, the estimated average equations for the learning curves for each keyboard were:

$$(2)\ CWPM_{Alphabetical} = 5.395594 + 0.226239*Time - 0.005125*Time^2$$

(3) $CWPM_{Predictive} = 6.024792 + 0.251622*Time - 0.005125*Time^2$

(4) $CWPM_{Hybrid} = 6.008488 + 0.258276*Time - 0.005125*Time^2$

Correlation within participant was captured via an auto regressive model with lag 1 (AR1). Indicator variables were chosen for keyboard as the learning curves obtained are specific to the levels selected. Keyboard had three levels modeled by using two indicator variables, $Keyboard_P$ and $Keyboard_H$, for the Predictive and Hybrid keyboards, respectively. The coefficients for these variables were compared to the baseline Alphabetical keyboard. Age, gender, and experience did not have a significant effect on learning.

The response versus fitted value plot appears in Figure 6-33. The plot shows a good model fit to the skill acquisition data. The standard deviations for the random effect terms are in Table 6-59. The combined standard deviations of the intercept, session, and keyboard random effects explained a large portion of the variability in the observed data when compared to the standard deviation of the error residual term. This implies that the variation between participants under similar conditions was greater compared to the within participant variation, as captured by the standard deviation of the residual error. Furthermore, there is evidence that inter-participant standard variations differed across keyboards, which implies that variability in performance among participants may be keyboard dependent. Specifically, inter-participant standard deviations were lower for typing with the Alphabetical than the nonstandard keyboards (see Appendix H). There was a moderate correlation between session and Predictive keyboard effects (0.585) and

Hybrid keyboard effects (0.504). Moreover, a large correlation (0.976) was observed between the Predictive and Hybrid keyboard effects.

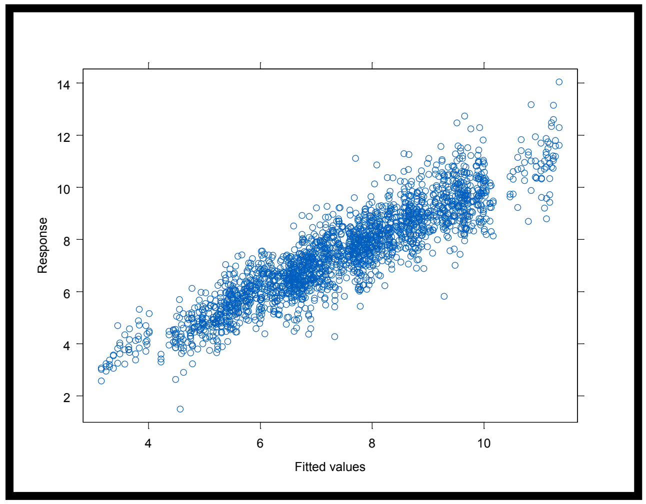


Figure 6-33 Response vs. Fitted Value Plot of the Quadratic Model for Text Entry Rate across Time

Parameter	***Std Dev***	***Correlation Matrix***		
		Intercept	**Session**	**Keyboard$_P$**
Intercept	1.137510			
Session	0.035870	0.086		
Keyboard$_P$	0.446107	0.229	0.585	
Keyboard$_H$	0.420167	0.425	0.504	0.976
Residual	0.865170			

Table 6-59 Estimates of the Variance Components of the Random Effect Terms of the Quadratic Model for Text Entry Rate across Time

Table 6-60 shows the model parameter estimates and the associated *p*-values. The average initial entry rates for keyboard layouts studied were 5.40 for the Alphabetical, 6.02 for the Predictive, and 6.00 for the Hybrid. Among the main effects, the coefficients for Keyboard$_P$ ($t(2142)= 4.177$, $p< .001$) and Keyboard$_H$ ($t(2142)= 4.289$, $p< .001$) were positive and highly significant. However, initial entry rates did not differ between the Hybrid and Predictive keyboards ($F(1,2142)= 0.032$, $p= 0.858$). The results indicated that average initial entry rates when typing with the nonstandard keyboards were generally higher than the average initial entry rates when typing the Alphabetical keyboard.

Parameter	*Value*	*Std. Error*	*DF*	*t-value*	*p-value*
Intercept	5.395594	0.332950	2142	16.20541	<.0001
Time	0.226239	0.014760	2142	4.17676	<.0001
$Time^2$	-0.005125	0.000499	2142	4.28964	<.0001
$Keyboard_P$	0.629198	0.150642	2142	15.32744	<.0001
$Keyboard_H$	0.612894	0.142878	2142	-10.51902	<.0001
$Keyboard_P$*Time	0.025383	0.006518	2142	3.89378	0.0001
$Keyboard_H$*Time	0.032037	0.006274	2142	5.10624	<.0001

Table 6-60 Parameter Estimates of the Quadratic Model using Actual Time

The average changes in entry rates are captured by the model parameters for the keyboard by time interactions. Trend in time data indicated a significant difference in learning rates between the Predictive and Alphabetical keyboards ($t\ (2142) = 3.894$, $p < .001$) and between the Hybrid and Alphabetical keyboards ($t\ (2142) = 5.106$, $p < .001$). The trend in time data did not indicate a significant difference in learning rates between the Predictive and Hybrid keyboards ($F(1,2142)=0.882$, $p= .348$).

As in the analysis using session, the model results indicated no crossover point for text entry rate because the Alphabetical keyboard was again estimated as being inferior to the Hybrid and Predictive keyboards. Also, there was no significant difference between the nonstandard keyboards, a finding supported by the positive parameter estimates for Predictive and Hybrid keyboards and expressed in Figure 6-34). The inference from this model is essentially the same as that from the model using session data, which indicates that the unequally spaced sessions did not affect the inference from the observed data.

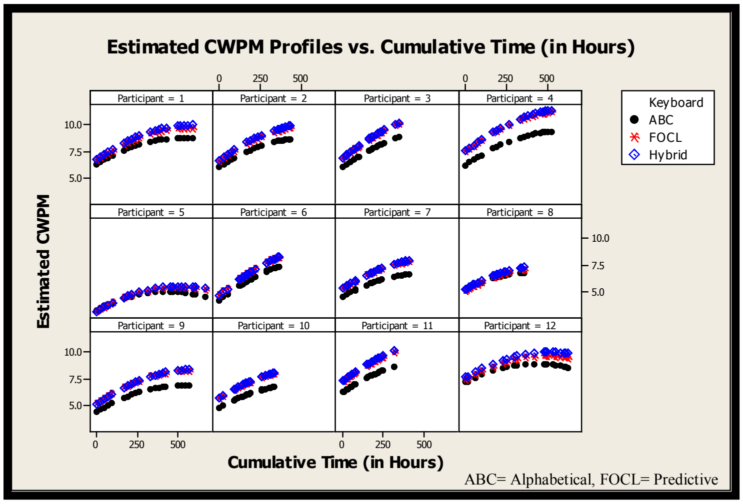


Figure 6-34 Quadratic Model Estimated Learning Trends for Text Entry Rates across Time

6.4.2.1.6 Predictive vs. Hybrid

It was anticipated that the Hybrid keyboard would outperform the Predictive keyboard, given the expected reduction of KSPC and visual search demands. However, analysis of initial rates and trend in time data did not reveal a significant difference in performance between the Predictive and Hybrid keyboards. In post-study interviews, participants indicated that deciding on the navigation path to the intended letter required more effort when typing with the Hybrid keyboard over the Predictive and Alphabetical keyboards. This suggested that an unexpected additional cognitive demand may have impeded Hybrid text entry performance. To explore this implication, longitudinal data for Reaction Time (measured as time between keystrokes), across all keyboards, was evaluated.

The mean reaction time (RT) across all twenty sessions were 0.214 (SD= 0.04) for Alphabetical, 0.264 (SD= 0.07) for Predictive, and 0.295 (SD= 0.07) for Hybrid. At the last session, mean reaction times were 0.191 (SD= 0.02) for Alphabetical, 0.231 (SD=

.04) for Predictive, and 0.259 (SD= 0.05) for Hybrid. Tables 6-61 and 6-62 provide descriptive statistics for RT by keyboard layout across all sessions and at session 20, respectively. Figure 6-35 depicts the mean reaction times by session and keyboard layout. Figure 6-36 shows the relationship between reaction times and session across keyboards for each participant.

Variable	*Keyboard*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
RT_{All}	Alphabetical	720	0.2139	0.0423	0.1529	0.1836	0.2074	0.2357	0.7056
	Predictive	720	0.2637	0.0661	0.1702	0.2142	0.2465	0.2980	0.5601
	Hybrid	720	0.2955	0.0740	0.1835	0.2378	0.2750	0..3396	0.6053

Table 6-61 Descriptive Statistics for Reaction Time across all Sessions

Variable	*Keyboard*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
RT_{20}	Alphabetical	36	0.1915	0.0239	0.1552	0.1692	0.1889	0.2148	0.2387
	Predictive	36	0.2310	0.0453	0.1702	0.1798	0.2133	0.2593	0.3604
	Hybrid	36	0.2587	0.0538	0.1883	0.2208	0.2356	0.2986	0.3670

Table 6-62 Descriptive Statistics for Reaction Times at Session 20

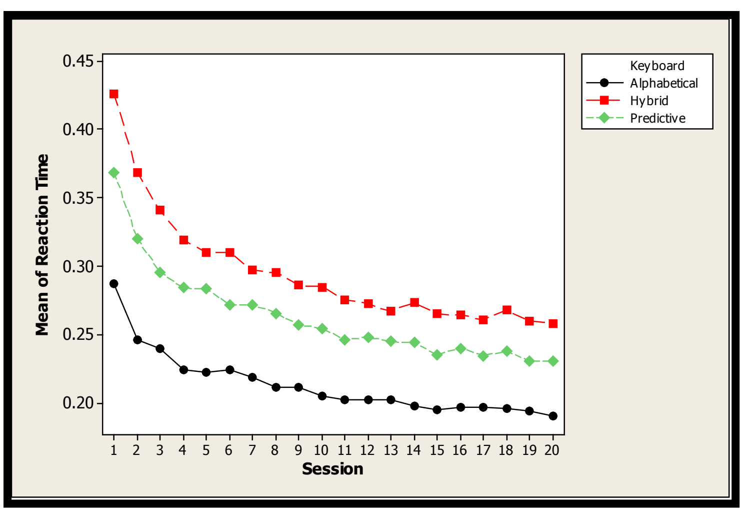


Figure 6-35 Mean Reaction Time for the Alphabetical, Predictive, and Hybrid Keyboards across Sessions

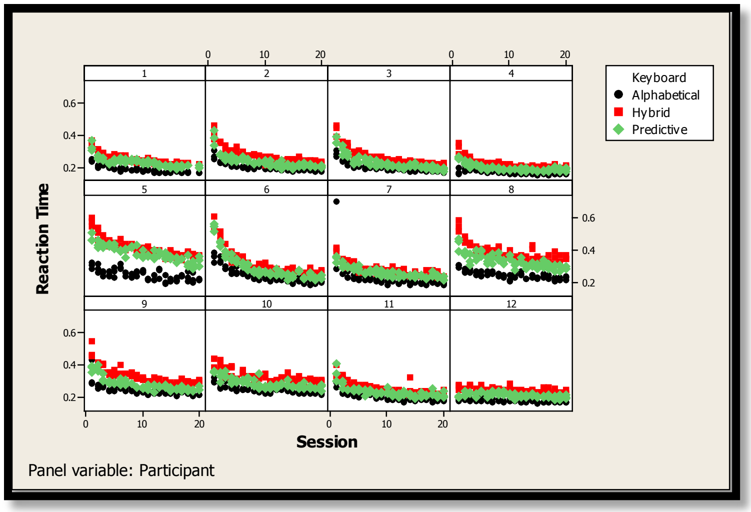


Figure 6-36 Participants' Reaction Times over Sessions

Among the models considered, the best fitting model to the data was a quadratic function. For this data, there was a non-linear relationship between keyboard reaction time and session. To minimize the non-linearity and to normalize the residuals, the inverse of reaction time was used. The best model was selected using a forward selection procedure and by inspection of associated residual plots and goodness-of-fit criteria. The following mixed effects quadratic model was selected for reaction time growth rates:

(1) 1/Reaction Time= 3.795788 - 0.816930*Keyboard$_P$ -1.243045*Keyboard$_H$ + 0.157379*Session - 0.004372*Session^2

Therefore, the estimated average equations for changes in reaction time for each keyboard were:

(2) $1/\text{Reaction Time}_{\text{Alphabetical}} = 3.795788 + 0.157379*\text{Session} - 0.004372*\text{Session}^2$

(3) $1/\text{Reaction Time}_{\text{Predictive}} = 2.978858 + 0.157379*\text{Session} - 0.004372*\text{Session}^2$

(4) $1/\text{Reaction Time}_{\text{Hybrid}} = 2.552743 + 0.157379*\text{Session} - 0.004372*\text{Session}^2$

Correlation within participant was captured via an auto regressive model with lag 1 (AR1). Indicator variables were chosen for keyboard as the learning curves obtained are specific to the levels selected. Keyboard had three levels modeled by using two indicator variables, Keyboard_P and Keyboard_H, for the Predictive and Hybrid keyboards, respectively. The coefficients for these variables were compared to the baseline Alphabetical keyboard. Age, gender, and experience did not have a significant effect on changes in reaction time.

The response versus fitted value plot appears in Figure 6-37. The plot shows a good model fit to the reaction time data. The standard deviations for the random effect terms are in Table 6-63. The standard deviation of the intercept random effect term explained a large portion of the variability in the observed data when compared to the standard deviation of the error residual term. Furthermore, there was evidence that inter-participant standard variations differed across keyboards, which implies that variability in performance among participants may be keyboard dependent. Specifically, inter-participant standard deviations were lower with the Hybrid keyboard than with the Predictive and Alphabetical keyboards (see Appendix H). Table 6-63 also shows that there was a moderate correlation between session and Predictive keyboard effects (0.496) and Hybrid keyboard effects (0.676). There was a high correlation between session^2 and session (-0.995) and between session^2 and the Hybrid keyboard (-0.729). There was also a high correlation (0.854) between the Predictive and Hybrid keyboard effects.

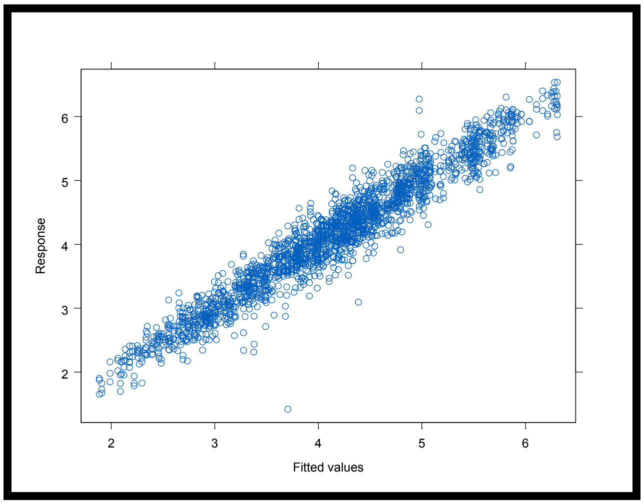


Figure 6-37 Response vs. Fitted Value Plot of the Model for Reaction Time

Parameter	***Std Dev***	***Correlation Matrix***			
		Intercept	**Session**	**Session^2**	**Keyboard$_P$**
Intercept	0.739627				
Session	0.066827	-0.463			
Session^2	0.002208	0.452	-0.995		
Keyboard$_P$	0.258709	-0.145	0.496	-0.557	
Keyboard$_H$	0.215388	-0.483	0.676	-0.729	0.854
Residual	0.255005				

Table 6-63 Estimates of the Variance Components of the Random Effect Terms of the Model for Reaction Time

Table 6-64 shows the model parameter estimates and the associated *p*-values. The average initial reaction times for the keyboard layouts studied were 0.263 for the Alphabetical, 0.336 for the Predictive, and 0.392 for the Hybrid. Among the main effects, the coefficients for Keyboard$_P$ ($t(2133)= -10.702$, $p< .001$) and Keyboard$_H$ ($t(2133)= -19.465$, $p< .001$) were negative and highly significant. Furthermore, initial reaction times significantly differed between the Hybrid and Predictive keyboards ($F(1,2133)= 106.729$, $p<.0001$). This means that average initial reaction time when typing with the Hybrid keyboard was generally greater than the average initial reaction times when typing with the Predictive and Alphabetical keyboards.

Parameter	*Value*	*Std. Error*	*DF*	*t-value*	*p-value*
Intercept	3.795788	0.214867	2133	17.666	<.0001
Session	0.157379	0.019875	2133	7.918	<.0001
$Session^2$	-0.004372	0.000675	2133	-6.478	<.0001
$Keyboard_P$	-0.816930	0.079335	2133	-10.702	<.0001
$Keyboard_H$	-1.243045	0.638589	2133	-19.465	<.0001

Table 6-64 Parameter Estimates of the Model for Reaction Time

Figure 6-38 presents the estimated profiles based on the fitted model. Learning rates are not captured in the parameters for this model since there was insufficient evidence to support a keyboard by session interaction. Although reaction times improved in time, the progression was not keyboard dependent. The model results did not indicate a crossover point, specifically because initially the Hybrid keyboard was inferior to the Alphabetical and Predictive keyboards. Unexpected cognitive demands seem to have impeded text entry performance for the Hybrid keyboard, which is likely the main reason why the Hybrid keyboard did not reach the expected entry rates.

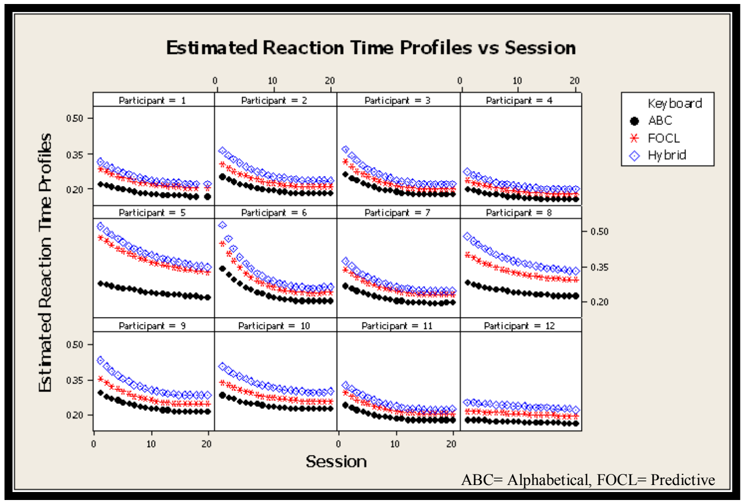


Figure 6-38 Estimated Learning Trends for Reaction Time

6.4.3 Error Rates

6.4.3.1 Uncorrected Error Rate

As in the previous experiments, accuracy was high across the keyboards. The mean uncorrected error rates across all twenty sessions were 0.91% (SD= 0.021) for Alphabetical, 0.59% (SD= 0.013) for Predictive, and 0.74% (SD= 0.015) for the Hybrid keyboard. The minimum (best) session averages were 0.50% for Alphabetical at session 16, 0.20% at session 17 for Predictive, and 0.27% at session 17 for Hybrid (see Figure 6-39). Table 6-65 provides descriptive statistics for uncorrected error rates by keyboard layout.

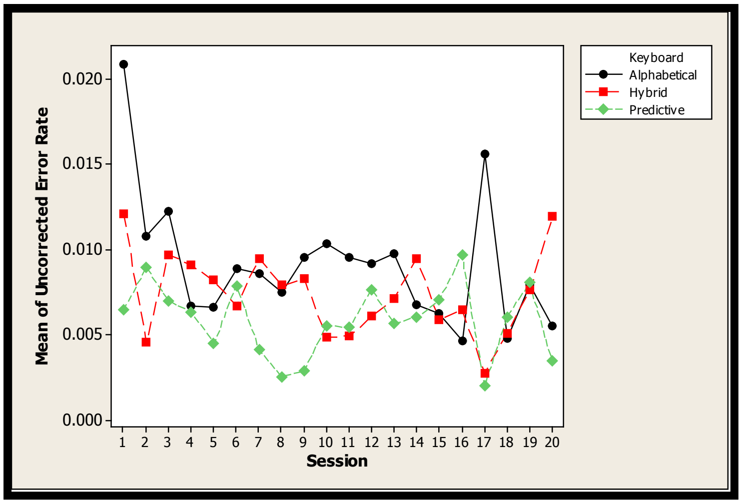


Figure 6-39 Uncorrected Error Rate over Sessions for the Alphabetical, Predictive, and Hybrid Keyboards

Variable	*Keyboard*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
UER	Alphabetical	717	.0091	.021	0	0	0	.010	.206
	Predictive	717	.0059	.013	0	0	0	.009	.119
	Hybrid	717	.0074	.015	0	0	0	.010	.137

Table 6-65 Descriptive Statistics for Uncorrected Errors Rate by Keyboard

As shown in Figure 6-40, the uncorrected error rates data across sessions was unstructured; highly zero inflated and with a few outliers. Given the lack of structure in the data, it was not possible to develop a mixed effects model for this variable. As an alternative analysis, logistic regression incorporating repeated measures was attempted to evaluate any keyboard layout or session effects. The data, however, lacked adequate diversity to explore learning effects across all sessions. As a consequence, the evaluation focused only on the differences in uncorrected error rates at the start (Session 1), midpoint (Session 10), and end (Session 20) of the longitudinal study. For the analysis, the uncorrected error rate data was coded to binary outcomes (assigned 0 for no errors occurring within a trial, and 1 for one or more errors occurring within a trial). In total, there were 1420 trials in error out of 2151 total trials.

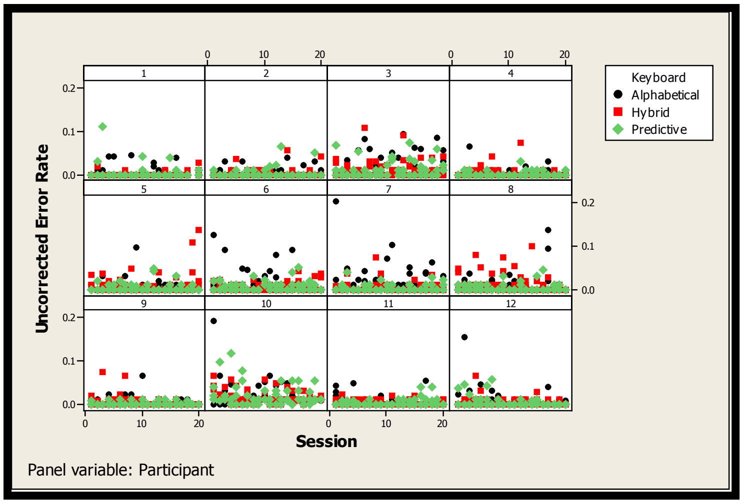


Figure 6-40 Individual Values of Keyboard Uncorrected Error Rates across Session for each Participant

The final model for the analysis was determined via forward selection procedures. The final model for this design included keyboard, session, age, a keyboard by session interaction, and an age by session interaction. Results indicated that a keyboard by session interaction ($X^2(4) = 22.970$, $p <.001$) was associated with uncorrected error rates (see Table 6-66). The results suggested that participants had fewer uncorrected errors by the end of the study, in comparison to Session 1, with both the Alphabetical (p =.091) and the Predictive (p =.025) keyboards. However, the Hybrid keyboard had a significantly lower uncorrected error rate at Session 10 than at Session 1 (p< .001) and at Session 20 (p< .001), see Figure 6-41 and Table 6-67.

Source	*Wald Chi-Square*	*df*	*Sig.*
Intercept	16.897	1	.000
Keyboard	8.299	2	.016
Session	12.446	2	.002
Age	5.317	1	.021
Session*Age	25.233	2	.000
Session*Keyboard	22.970	4	.000

Table 6-66 Tests of Model Effects for Uncorrected Error Rate

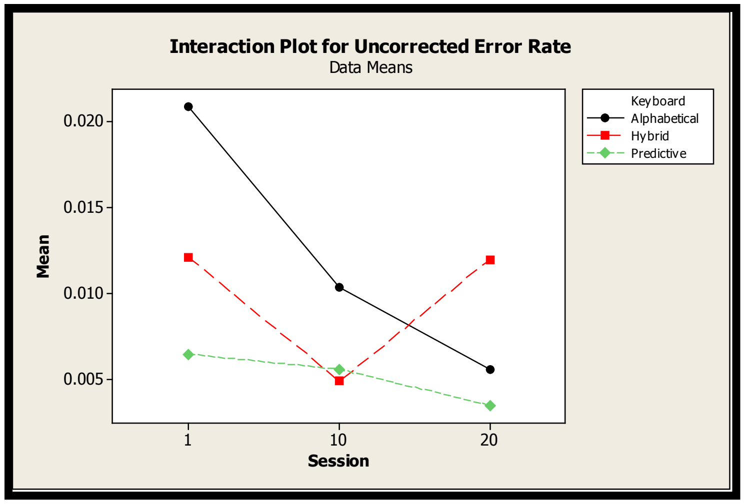


Figure 6-41 Keyboard by Session Interaction Plot for Uncorrected Error Rate

Keyboard	*(I) Session*	*(J) Session*	*Mean Difference (I-J)*	*Std. Error*	*df*	*Sig.*	*95% Wald Confidence Interval for Difference*	
							Lower	*Upper*
Alphabetical	1.00	10.00	.08	.051	1	.129	-.02	.18
		20.00	.13	.074	1	.091	-.02	.27
	10.00	1.00	-.08	.051	1	.129	-.18	.02
		20.00	.05	.052	1	.361	-.05	.15
	20.00	1.00	-.13	.074	1	.091	-.27	.02
		10.00	-.05	.052	1	.361	-.15	.05
Predictive	1.00	10.00	-.03	.031	1	.285	-.09	.03
		20.00	-.05	.024	1	.025	-.10	.00
	10.00	1.00	.03	.031	1	.285	-.03	.09
		20.00	-.02	.038	1	.590	-.10	.05
	20.00	1.00	.05	.024	1	.025	.01	.10
		10.00	.02	.038	1	.590	-.05	.10
Hybrid	1.00	10.00	.20	.044	1	.000	.11	.29
		20.00	.03	.029	1	.306	-.03	.09
	10.00	1.00	-.20	.044	1	.000	-.29	-.11
		20.00	-.17	.041	1	.000	-.25	-.09
	20.00	1.00	-.03	.029	1	.306	-.09	.03
		10.00	.17	.041	1	.000	.09	.25

Table 6-67 Keyboard by Session Pairwise Comparisons for Uncorrected Error Rate

It was hypothesized that expert uncorrected error rate for the Predictive keyboard would be greater than the error rates for the Alphabetical and Hybrid keyboards. However, results presented in Table 6-68 suggested that participants had greater uncorrected errors at the end of the study with the Hybrid keyboard than with the Predictive (p =.056) and the Alphabetical (p= .058) keyboards.

Session	(I) Keyboard	(J) Keyboard	Mean Difference (I-J)	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference	
							Lower	Upper
1	Alphabetical	Predictive	.17	.062	1	.006	.05	.29
		Hybrid	.00	.047	1	1.000	-.09	.09
	Predictive	Alphabetical	-.17	.062	1	.006	-.29	-.05
		Hybrid	-.17	.041	1	.000	-.25	-.09
	Hybrid	Alphabetical	.00	.047	1	1.000	-.09	.09
		Predictive	.17	.041	1	.000	.09	.25
10	Alphabetical	Predictive	.06	.039	1	.124	-.02	.14
		Hybrid	.12	.050	1	.014	.02	.22
	Predictive	Alphabetical	-.06	.039	1	.124	-.14	.02
		Hybrid	.06	.043	1	.134	-.02	.15
	Hybrid	Alphabetical	-.12	.050	1	.014	-.22	-.02
		Predictive	-.06	.043	1	.134	-.15	.02
20	Alphabetical	Predictive	.00	.019	1	.652	-.05	.03
		Hybrid	-.09	.050	1	.058	-.19	.00
	Predictive	Alphabetical	.01	.019	1	.652	-.03	.05
		Hybrid	-.09	.045	1	.056	-.17	.00
	Hybrid	Alphabetical	.09	.050	1	.058	.00	.19
		Predictive	.09	.045	1	.056	.00	.17

Table 6-68 Session by Keyboard Pairwise Comparisons for Uncorrected Error Rate

Results also indicated that an age by session interaction ($X^2(2) = 25.233$, p <.001) was associated with the uncorrected error rates (see Figure 6-42), specifically that participants less than 40 years old had fewer uncorrected errors at Session 10 than at Session 1 (p= .003) and at Session 20 (p <.001, see Table 6-69). The data suggested that participants 40 or more years old had a lower uncorrected error rate at Session 20 than at Session 1 (p = .037). It was unexpected for session to have an effect on uncorrected error rates, as expertise generally does not reduce uncorrected errors. These findings represent uncorrected error rates at discrete points in time and do not reflect trends in time as function of learning.

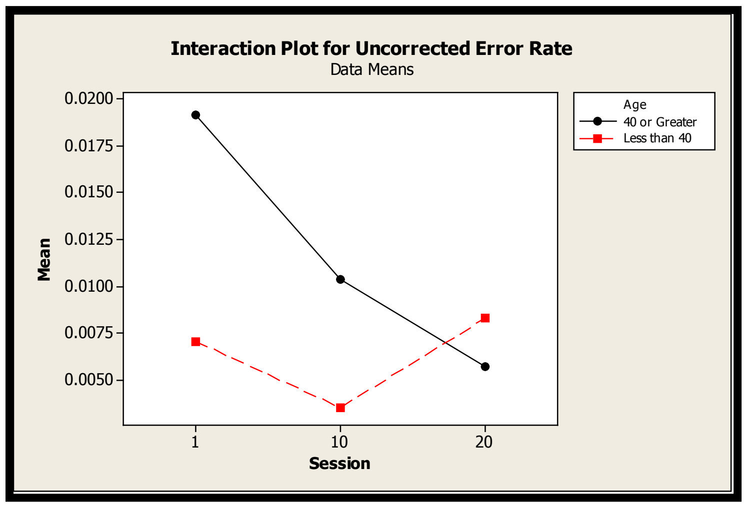


Figure 6-42 Age by Session Interaction Plot for Uncorrected Error Rate

Age	(I) Session	(J) Session	Mean Difference (I-J)	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference	
							Lower	Upper
Less	1.00	10.00	.12	.040	1	.003	.04	.20
		20.00	-.03	.037	1	.454	-.10	.05
	10.00	1.00	-.12	.040	1	.003	-.20	-.04
		20.00	-.15	.021	1	.000	-.19	-.10
	20.00	1.00	.03	.037	1	.454	-.05	.10
		10.00	.15	.021	1	.000	.10	.19
More	1.00	10.00	.03	.026	1	.202	-.02	.09
		20.00	.09	.045	1	.037	.01	.18
	10.00	1.00	-.03	.026	1	.202	-.09	.02
		20.00	.06	.037	1	.100	-.01	.13
	20.00	1.00	-.09	.045	1	.037	-.18	.00
		10.00	-.06	.037	1	.100	-.13	.01

Table 6-69 Age by Session Pairwise Comparisons for Uncorrected Error Rate

6.4.3.2 Total Error Rate

The mean total error rates (TER), across all twenty sessions, were 5.0% (SD= 0.055) for Alphabetical, 4.5% (SD= 0.049) for Predictive, and 4.7% (SD= 0.048) for the Hybrid keyboard. The minimum (best) session averages were 3.8 % for Alphabetical at session 14, 3.3% at session 10 for Predictive, and 3.4% at session 13 for Hybrid (see Figure 6-43). Table 6-70 provides descriptive statistics for TER by keyboard layout.

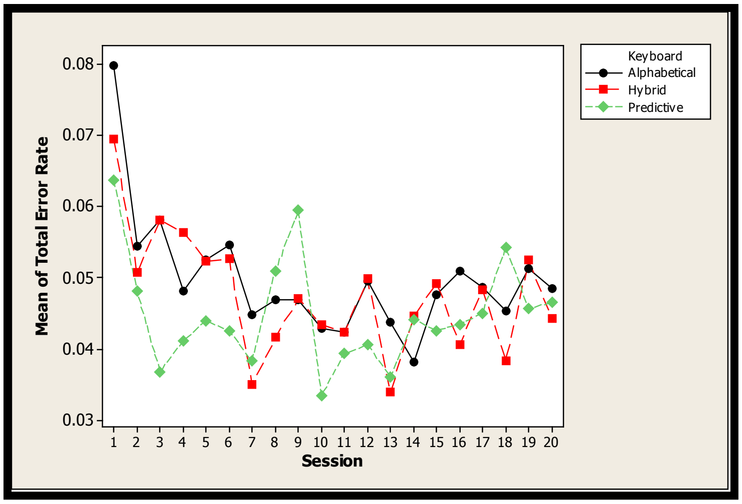


Figure 6-43 Total Error Rate over Sessions for the Alphabetical, Predictive, and Hybrid Keyboards

Variable	*Keyboard*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
TER	Alphabetical	717	.0498	.055	0	.018	.036	.061	.331
	Predictive	717	.0448	.049	0	.011	.030	.061	.316
	Hybrid	717	.0475	.048	0	.011	.036	.068	.292

Table 6-70 Descriptive Statistics for Total Error Rate by Keyboard

Similar to the uncorrected error rates, the trends in time for the total error rates lacked the structure needed to develop a mixed effects model (see Figure 6-44). As an alternative analysis, logistic regression incorporating repeated measures was attempted to evaluate any keyboard layout or session effects. However, the data also lacked adequate diversity to explore learning effects across all sessions. As a consequence, only the differences in TER at the start (Session 1), midpoint (Session 10), and end (Session 20) of the study were evaluated. The TER data was coded to binary outcomes (assigned 0 for no errors occurring within a trial, and 1 for one or more errors occurring within a trial). In total, there were 1814 trials in error out of 2151 total trials.

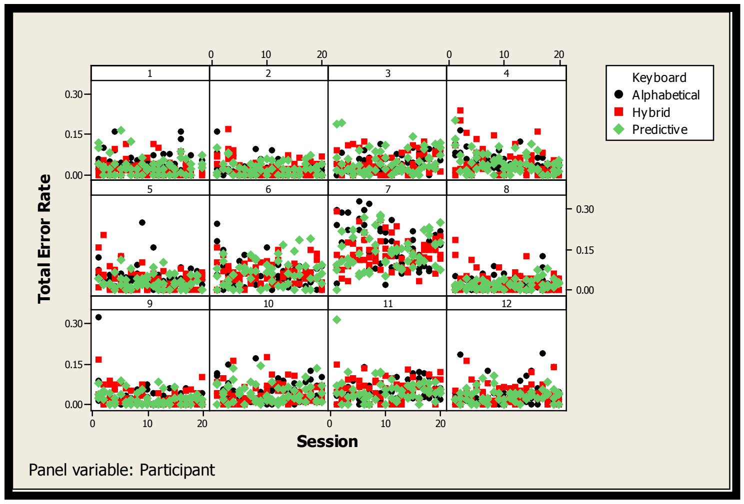


Figure 6-44 Individual Values of Keyboard Total Error Rate across Session for each Participant

The final model for this design included keyboard, session, age, and an age by session interaction, as determined via forward selection procedures. Results indicated that keyboard layout ($X^2(2) = 26481$, $p <.001$) was associated with total error rates (see Table 6-71). Specifically, participants made fewer mistakes overall with the Hybrid keyboard than with the Alphabetical ($p <.001$, see Table 6-72).

Source	*Wald Chi-Square*	*df*	*Sig.*
Intercept	124.582	1	.000
Keyboard	26.481	2	.000
Session	3.910	2	.142
Age	4.233	1	.040
Session*Age	36.662	2	.000

Table 6-71 Tests of Model Effects for Total Error Rate

(I) Keyboard	(J) Keyboard	Mean Difference (I-J)	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference	
						Lower	Upper
Alphabetical	Predictive	.04	.018	1	.015	.01	.08
	Hybrid	.09	.017	1	.000	.05	.12
Predictive	Alphabetical	-.04	.018	1	.015	-.08	.00
	Hybrid	.04	.019	1	.025	.01	.08
Hybrid	Alphabetical	-.09	.017	1	.000	-.12	-.05
	Predictive	-.04	.019	1	.025	-.08	.00

Table 6-72 Keyboard by Session Pairwise Comparisons for Total Error Rate

Results also indicated that an age by session interaction ($X^2(2) = 36.662$, $p< .001$) was associated with the TER (see Figure 6-45), specifically that participants less than 40 years old made significantly more mistakes at Session 1 than at Session 10 ($p< .001$) and at Session 20 ($p <.001$, see Table 6-73). Similarly, participants 40 or more years old made more mistakes at Session 1 than at Session 10 ($p= .003$) and at Session 20 ($p <.001$). Younger participants made fewer mistakes than older participants at Session 10 ($p< .001$) and at Session 20 ($p= .002$, see Table 6-74).

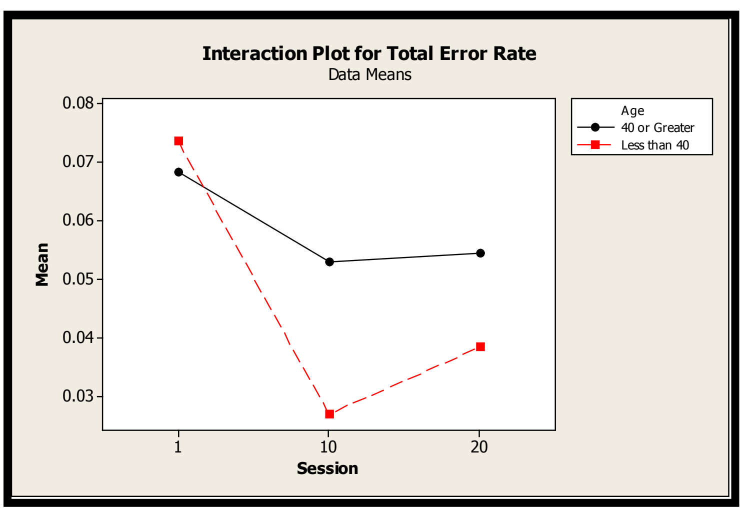


Figure 6-45 Age by Session Interaction Plot for Total Error Rate

Age	(I) Session	(J) Session	Mean Difference (I-J)	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference	
							Lower	Upper
Less than 40	1	10	.08	.016	1	.000	.05	.12
		20	.10	.027	1	.000	.05	.15
	10	1	-.08	.016	1	.000	-.12	-.05
		20	.01	.016	1	.379	-.02	.05
	20	1	-.10	.027	1	.000	-.15	-.05
		10	-.01	.016	1	.379	-.05	.02
40 or more	1	10	-.05	.017	1	.003	-.08	-.02
		20	-.04	.010	1	.000	-.06	-.02
	10	1	.05	.017	1	.003	.02	.08
		20	.01	.021	1	.564	-.03	.05
	20	1	.04	.010	1	.000	.02	.06
		10	-.01	.021	1	.564	-.05	.03

Table 6-73 Age by Session Pairwise Comparisons for Total Error Rate

Session	(I) Age	(J) Age	Mean Difference (I-J)	Std. Error	df	Sig.	95% Wald Confidence Interval for Difference	
							Lower	Upper
1	<40	≥40	.02	.029	1	.420	-.03	.08
	≥40	<40	-.02	.029	1	.420	-.08	.03
10	<40	≥40	-.11	.035	1	.001	-.18	-.04
	≥40	<40	.11	.035	1	.001	.04	.18
20	<40	≥40	-.11	.037	1	.002	-.19	-.04
	≥40	<40	.11	.037	1	.002	.04	.19

Table 6-74 Session by Age Pairwise Comparisons for Total Error Rate

6.4.4 Efficiency

6.4.4.1 KSPC

The mean KSPC across all 20 sessions were 8.506 (SD= 1.14) for Alphabetical, 6.231 (SD= 0.75) for Predictive, and 5.534 (SD= 0.73) for Hybrid. At the last session, mean KSPC were 8.278 (SD= 1.00) for Alphabetical, 6.120 (SD= 0.73) for Predictive, and 5.423 (SD= 0.80) for Hybrid. Tables 6-75 and 6-76 provide descriptive statistics for mean KSPC by keyboard layout across all sessions and at Session 20, respectively. Figure 6-46 shows mean KSPC by session and keyboard layout. Figure 6-47 shows the relationship between entry rates and session, across keyboards for each participant.

Variable	*Keyboard*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
$KSPC_{All}$	Alphabetical	720	8.506	1.138	6.629	7.789	8.245	8.849	13.644
	Predictive	720	6.231	0.755	4.247	5.731	6.200	6.685	10.560
	Hybrid	720	5.534	0.726	3.828	5.030	5.400	5.928	8.185

Table 6-75 Descriptive Statistics for KSPC across all Sessions

Variable	*Keyboard*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
$KSPC_{20}$	Alphabetical	36	8.278	1.004	7.048	7.504	7.962	8.622	10.881
	Predictive	36	6.120	0.727	4.600	5.507	6.228	6.721	7.763
	Hybrid	36	5.423	0.804	4.011	4.940	5.288	6.009	7.045

Table 6-76 Descriptive Statistics for KSPC at Session 20

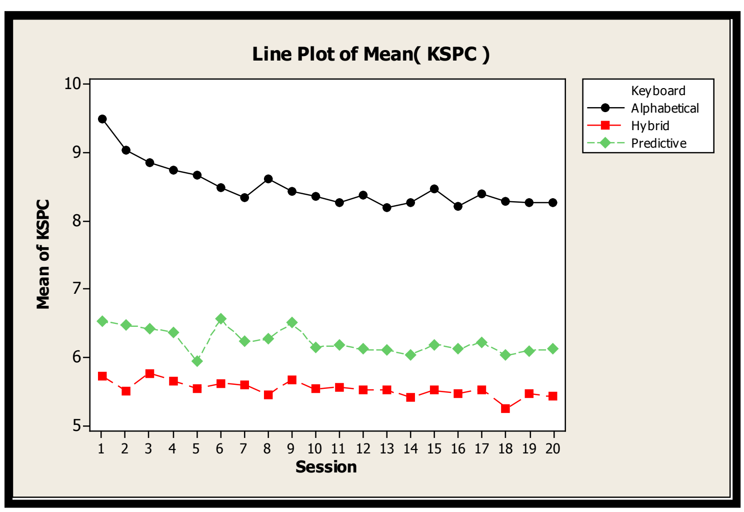


Figure 6-46 KSPC for the Alphabetical, Predictive, and Hybrid Keyboards across Sessions

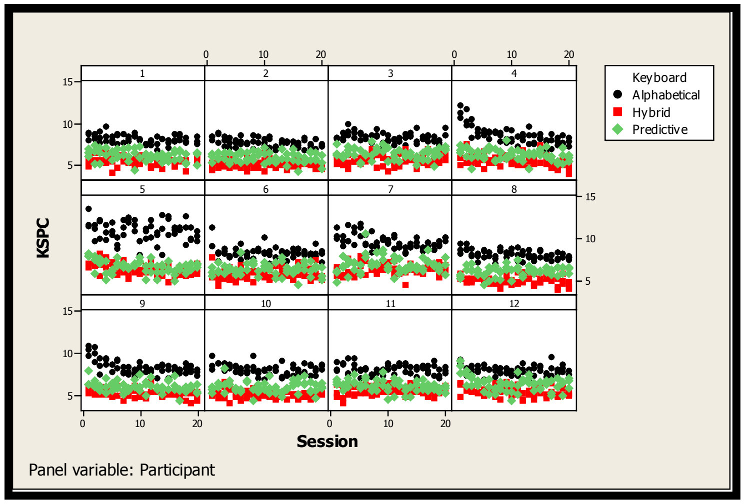


Figure 6-47 Participants' KSPC over Sessions

A mixed effects model was employed for the analysis of the KSPC longitudinal data. The best fitting model to the data, among the models considered, was a quadratic model. The best model was selected using a forward selection procedure and by inspection of associated residual plots and goodness-of-fit criteria. The following mixed effects quadratic model was selected for KSPC learning effects:

(1) $KSPC = 9.234285 + 0.254239*Gender - 2.834092*Keyboard_P - 3.680600*Keyboard_H - 0.153137*Session + 0.109459*Keyboard_P*Session + 0.139765*Keyboard_H*Session + 0.005231*Session^2 - 0.004097*Keyboard_P*Session^2 - 0.005280*Keyboard_H*Session^2$

Therefore, the estimated average equations for the KSPC learning curves for each keyboard were:

(2) $KSPC_{Alphabetical} = 9.234285 + 0.254239*Gender - 0.153137*Session + 0.005231*Session^2$

(3) $KSPC_{Predictive}$= 6.400193+ 0.254239*Gender -0.043678*Session + 0.001134*$Session^2$

(4) $KSPC_{Hybrid}$= 5.553685+ 0.254239*Gender – 0.013372*Session – 0.000049*$Session^2$

Correlation within participant was captured via an auto regressive model with lag 1 (AR1). Indicator variables were chosen for keyboard and gender as the learning curves obtained are specific to the levels selected. Keyboard had three levels modeled by using two indicator variables, $Keyboard_P$ and $Keyboard_H$, for the Predictive and Hybrid keyboards, respectively. The coefficients for these variables were compared to the baseline Alphabetical keyboard. Gender had two levels, with females coded as 0 and males coded as 1. The quadratic model incorporated the effect of keyboard, gender, and session on KSPC, as age and experience did not have a significant effect on KSPC progression.

The response versus fitted value plot appears in Figure 6-48. The plot shows an adequate model fit to the longitudinal KSPC data. The standard deviations for the random effect terms are in Table 6-77. Collectively, the standard deviations of the intercept, session, and keyboard random effects explained a large portion of the variability in the observed data when compared to the standard deviation of the error residual term. Furthermore, there is evidence that inter-participant standard variations differed across keyboards, which implies that variability in performance among participants may be keyboard dependent. Specifically, inter-participant standard deviations were lower with the Hybrid keyboard than with the Predictive and Alphabetical keyboards (see Appendix H). Table 6-77 also shows that there was a moderate correlation between session and Predictive keyboard effects (0.502) and Hybrid keyboard effects (0.571), and a high correlation (0.978) between the Predictive and Hybrid keyboard effects.

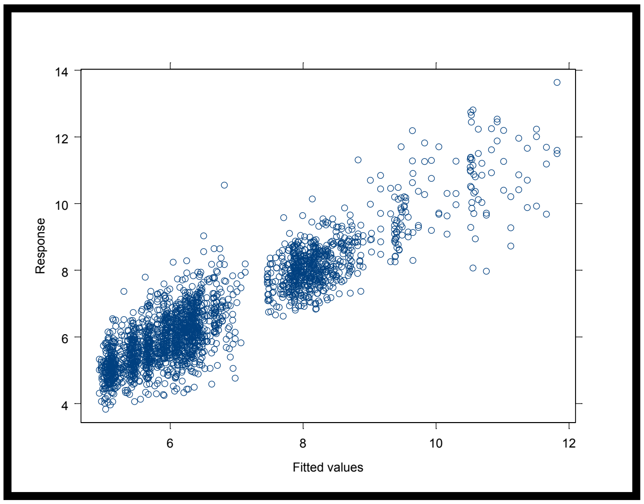


Figure 6-48 Response vs. Fitted Value Plot of the Model for KSPC

Parameter	*Std Dev*	*Correlation Matrix*		
		Intercept	**Session**	**Keyboard$_P$**
Intercept	0.957403			
Session	0.017389	-0.609		
Keyboard$_P$	0.703447	-0.974	0.502	
Keyboard$_H$	0.561756	-0.932	0.571	0.978
Residual	0.701130			

Table 6-77 Estimates of the Variance Components of the Random Effect Terms of the Model for KSPC

Table 6-78 shows the model parameter estimates and the associated *p*-values. Results indicated a marginally significant main effect for gender (t(10)= 2.32, p= .043), implying that initial KSPC differed by gender, with males initially requiring more keystrokes per character than females. The average initial KSPC for males were 9.49 for the Alphabetical, 6.65 for the Predictive, and 5.81 for the Hybrid. For females, the average initial KSPC were 9.23 for the Alphabetical, 6.40 for the Predictive, and 5.55 for the Hybrid.

Among the main effects, the coefficients for $Keyboard_P$ ($t(2129)= -12.003$, $p< .0001$) and $Keyboard_H$ ($t(2129)= -18.813$, $p< .0001$) were negative and highly significant. Furthermore, initial KSPC significantly differed between the Hybrid and Predictive keyboards ($F(1,2129)= 45.583$, $p< .0001$), which means that average initial KSPC when typing with the Hybrid keyboard was generally lower than the average initial entry rates when typing the Predictive and Alphabetical keyboards.

Parameter	Value	Std. Error	DF	t-value	p-value
Intercept	9.234285	0.293820	2129	31.42832	<.0001
Session	-0.153137	0.018975	2129	-8.07055	<.0001
Gender	0.254239	0.109595	10	2.31980	.0428
$Session^2$	0.005231	0.000848	2129	6.16863	<.0001
$Keyboard_P$	-2.383409	0.236110	2129	-12.00327	<.0001
$Keyboard_H$	-3.680600	0.195641	2129	-18.81301	<.0001
$Keyboard_P$* Session	0.109459	0.026452	2129	4.13790	<.0001
$Keyboard_H$*Session	0.139765	0.024039	2129	5.81396	<.0001
$Keyboard_P$* $Session^2$	-0.004097	0.001226	2129	-3.34212	.0008
$Keyboard_H$*$Session^2$	-0.005280	0.000111	2129	-4.73879	<.0001

Table 6-78 Parameter Estimates of the KSPC Quadratic Model using Session

Learning rates are captured in the model parameters via the keyboard by session interactions. Herein, learning rates were estimated by both linear and quadratic interactions. Linear trend in time data indicated a significant difference in KSPC learning rates between the Predictive and Alphabetical keyboards ($t (2129) =4.138$, $p< .0001$) and between the Hybrid and Alphabetical keyboards ($t (2129) =5.814$, $p< .0001$). However, linear trend in time data did not suggest a significant difference in learning rates between the Predictive and Hybrid keyboards ($F(1,2129)=1.511$, $p=.219$). Furthermore, quadratic trend in time data indicated a significant difference in KSPC learning rates between the

Predictive and Alphabetical keyboards (t (2129) = -3.342, p< .0001) and between the Hybrid and Alphabetical keyboards (t (2129) = -4.739, p< .0001). Although, quadratic trend in time data did not suggest a significant difference in learning rates between the Predictive and Hybrid keyboards (F(1,2129)=1.070, p=.301). This implies that although the Hybrid had better initial KSPC values, the Alphabetical had the greatest learning rate for the reduction in KSPC.

Figure 6-49 presents the estimated profiles based on the fitted model. KSPC rates decreased over time for all participants. Leveling off of KSPC is not apparent in the figure. Therefore, it is possible that greater reduction in KSPC for the keyboards may be attained with time, but the data are insufficient to identify an observed lower limit.

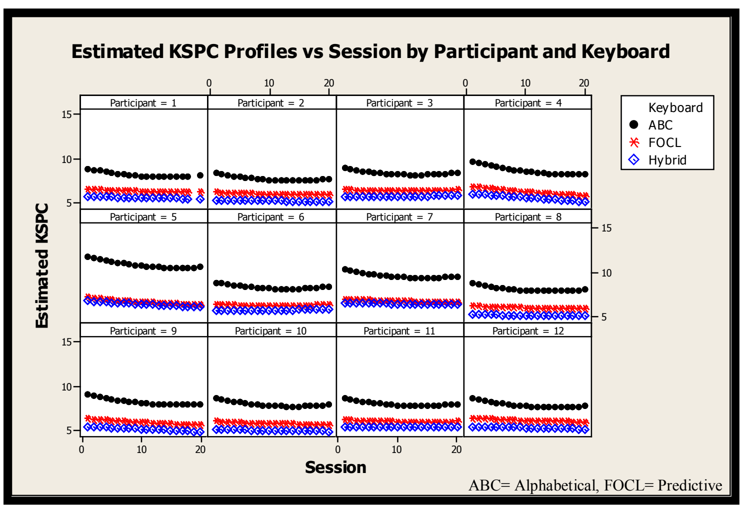


Figure 6-49 Estimated Learning Trends for KSPC

6.4.4.2 Movement Inefficiency

Selector movement inefficiency (MI), a measure of the percentage of keystrokes employed more than optimal, was also analyzed in this experiment. Mean selector movement inefficiency across all sessions were 16.8% more than optimal with the Alphabetical keyboard, 22.8% more with the Predictive, and 20.6% more with the Hybrid. Table 6-79 provides descriptive statistics for selector movement inefficiency by keyboard across all sessions. Figure 6-50 shows mean selector movement inefficiency by keyboard across all sessions.

Variable	*Keyboard*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
MI	Alphabetical	717	0.168	0.176	.000	0.070	0.111	0.189	1.748
	Predictive	716	0.228	0.105	.000	0.158	0.215	0.284	0.754
	Hybrid	716	0.206	0.119	.000	0.121	0.179	0.260	0.687

Table 6-79 Descriptive Statistics for Movement Inefficiency across all Sessions

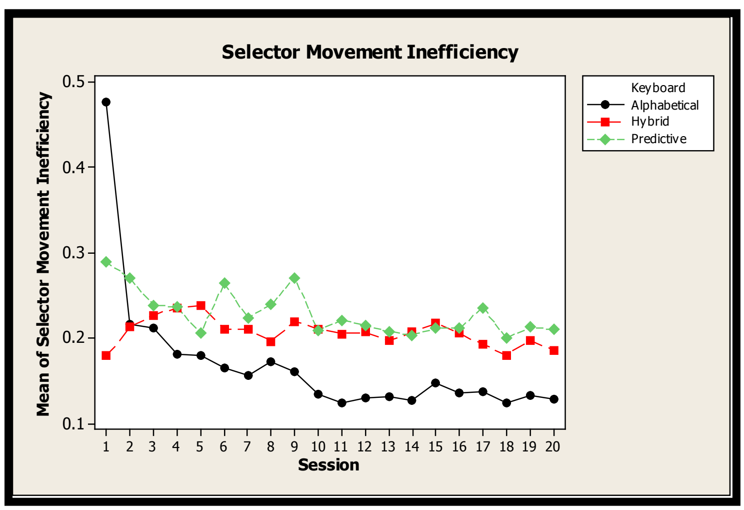


Figure 6-50 Movement Inefficiency for the Alphabetical, Predictive, and Hybrid Keyboards across Sessions

As in the previous experiments, there was a significant relationship between the selector movement inefficiency across the keyboards (see Table 6-80 and Figure 6-51). These relationships indicated that participants' levels of efficiency were consistent across all keyboards.

		Alphabetical	Predictive	Hybrid
Alphabetical	Pearson Correlation	1	.417**	.303**
	Sig. (2-tailed)		.000	.000
	N	717	716	716
Predictive	Pearson Correlation	.417**	1	.446**
	Sig. (2-tailed)	.000		.000
	N	716	716	715
Hybrid	Pearson Correlation	.303**	.446**	1
	Sig. (2-tailed)	.000	.000	
	N	716	715	716

**. Correlation is significant at the 0.01 level (2-tailed).

Table 6-80 Correlation between Movement Inefficiency and Keyboard

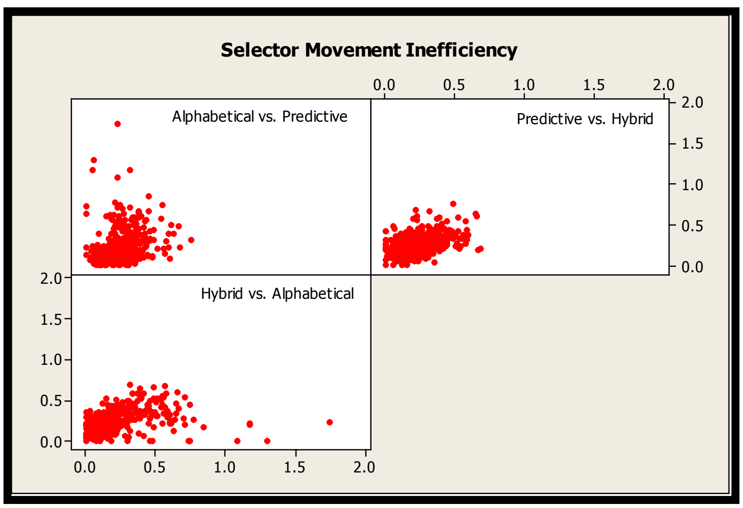


Figure 6-51 Scatter Plots of Keyboard Movement Inefficiency

For selector movement inefficiency, the evaluation focused only on the differences in efficiency at the start (Session 1), midpoint (Session 10), and end (Session 20) of the longitudinal study. Figure 6-52 shows mean selector movement inefficiency across keyboards at Session 1, 10 and 20.

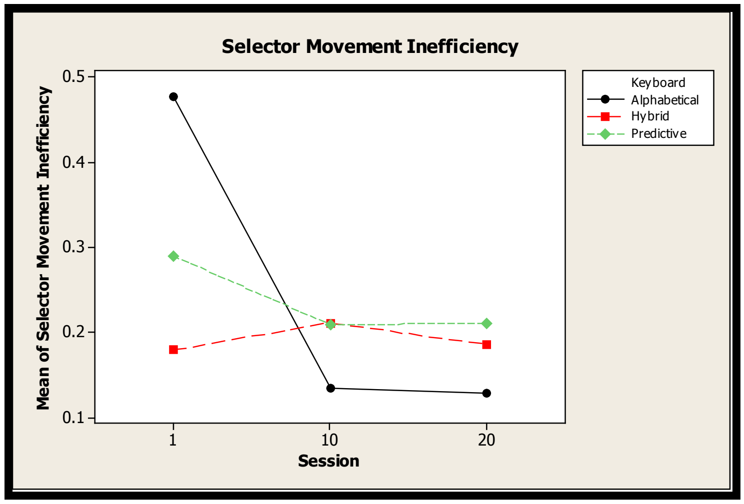


Figure 6-52 Movement Inefficiency for the Alphabetical, Predictive, and Hybrid Keyboards at Session 1, 10, and 20

The mean selector movement inefficiency across keyboards, for Session 1, was 47.7% for Alphabetical, 28.9% for Predictive, and 17.9% for Hybrid. Nonparametric analysis with a Friedman's test indicated that selector movement inefficiency for Session 1 was not different across keyboards ($X^2(2)= 4.50$, $p= .105$, see Appendix H). Table 6-81 provides descriptive statistics for movement inefficiency by keyboard at Session 1.

Variable	*Keyboard*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
MI	Alphabetical	36	0.477	0.417	0	0.144	0.361	0.743	1.749
	Predictive	36	0.289	0.172	0	0.181	0.256	0.442	0.677
	Hybrid	36	0.179	0.157	0	0.057	0.171	0.218	0.668

Table 6-81 Descriptive Statistics for Movement Inefficiency by Keyboard for Session 1

At Session 10, the mean selector movement inefficiency across keyboards was 13.4% for Alphabetical, 20.9% for Predictive, and 21.1% for Hybrid. Table 6-82 provides descriptive statistics for selector movement inefficiency by keyboard at Session 10.

Variable	Keyboard	N	M	SD	Min	Q1	Mdn	Q3	Max
MI	Alphabetical	36	0.134	0.096	0	0.078	0.111	0.157	0.366
	Predictive	36	0.209	0.095	0	0.135	0.191	0.276	0.441
	Hybrid	36	0.211	0.112	0	0.123	0.187	0.284	0.491

Table 6-82 Descriptive Statistics for Movement Inefficiency by Keyboard for Session 10

A Friedman's test indicated that selector movement inefficiency at Session 10 was different across keyboards ($X^2(2)= 15.17$, $p< .001$). Post-hoc analysis, based on Friedman rank-averages, indicated significantly greater selector movement efficiency with the Alphabetical over the Predictive ($p< .01$) and the Hybrid ($p< .01$) keyboards.

At Session 20, the mean selector movement inefficiencies across keyboards were 12.7% for Alphabetical, 21.0% for Predictive, and 18.5% for Hybrid. Table 6-83 provides descriptive statistics for selector movement inefficiency by keyboard.

Variable	Keyboard	N	M	SD	Min	Q1	Mdn	Q3	Max
MI	Alphabetical	36	0.127	0.122	0.022	0.051	0.080	0.152	0.494
	Predictive	36	0.210	0.088	0.084	0.130	0.204	0.280	0.426
	Hybrid	36	0.185	0.100	0.035	0.109	0.156	0.249	0.475

Table 6-83 Descriptive Statistics for Movement Inefficiency by Keyboard for Session 20

A Friedman's test indicated that selector movement inefficiency at Session 20 was significantly different across keyboards ($X^2(2)= 12.50$, $p =.002$). As in Session 10, post-hoc analysis, based on Friedman rank-averages, indicated that participants were significantly more efficient in typing with the Alphabetical over the Predictive ($p <.01$) and the Hybrid ($p< .01$) keyboards. However, the superior performance with the Alphabetical keyboard as indicated by this measure did not overcome the initial inefficiencies of the layout.

As explained previously, this measure is primarily indicative of participants' tendencies to overshoot an intended letter when using typematic keying. It was hypothesized that the Predictive keyboard would have the greatest selector movement inefficiencies across all sessions. This did not appear to be the case, given that there was no indication of a significant difference in selector movement efficiency between Predictive and Hybrid within Sessions 1, 10, or 20. With learning, participants were expected to become more efficient across all keyboards. Friedman tests showed no significant effect between sessions in selector movement efficiency, for Hybrid ($X^2(2)= 0.67$, $p =.716$) or Predictive ($X^2(2)= 4.50$, $p =.105$) keyboards. However, the Friedman tests indicate that selector movement inefficiency for the Alphabetical keyboard was significantly different across sessions ($X^2(2)= 13.17$, $p =.002$), with participants more efficient at Session 20 than at Session 1 ($p <.01$).

6.4.4.3 Typematic Keying

Typematic keying rates for Alphabetical, Predictive and Hybrid keyboards, across all sessions, comprised 59.1%, 39.3%, and 26.8% of the observed keystrokes, respectively. Table 6-84 provides descriptive statistics for TE by keyboard across all sessions.

Variable	*Keyboard*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
TE	Alphabetical	717	.591	.178	.000	.561	.655	.695	.777
	Predictive	716	.393	.167	.000	.291	.419	.527	.726
	Hybrid	716	.268	.132	.000	.174	.271	.375	.551

Table 6-84 Descriptive Statistics for Typematic Events across all Sessions

As in the previous experiments, there was a significant relationship between the rates of use of typematic keying across the keyboards (see Table 6-85 and Figure 6-53).

These relationships indicates that if a participant employed typematic keying with the any one keyboard, then it was likely that participant also employed typematic keying, with similar rates of use, with the other keyboards.

		Alphabetical	Predictive	Hybrid
Alphabetical	Pearson Correlation	1	.746**	.717**
	Sig. (2-tailed)		.000	.000
	N	717	716	716
Predictive	Pearson Correlation	.746**	1	.816**
	Sig. (2-tailed)	.000		.000
	N	716	716	715
Hybrid	Pearson Correlation	.717**	.816**	1
	Sig. (2-tailed)	.000	.000	
	N	716	715	716

**. Correlation is significant at the 0.01 level (2-tailed).

Table 6-85 Correlation between Typematic Events and Keyboard

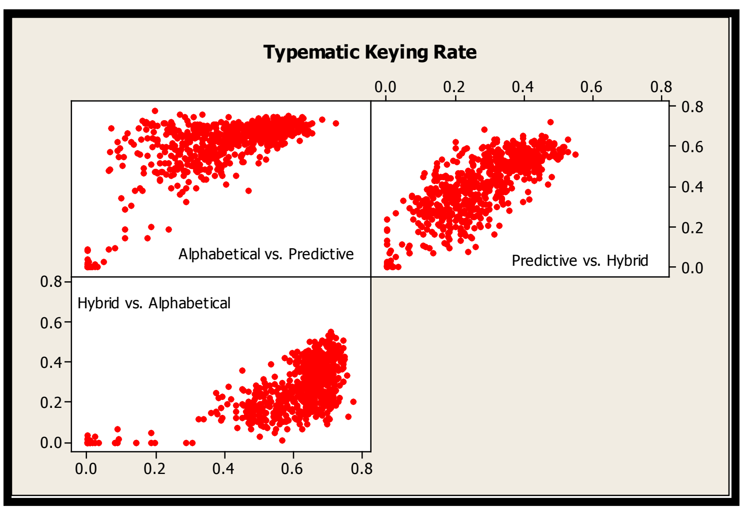


Figure 6-53 Scatter Plots of Keyboard Typematic Events

Figure 6-54 provides TE rate of use for each individual across sessions. Participant 5 and 7 showed very different behavior than the others in the study. In fact, participant 7 stopped using typematic keying after a few sessions, saying that the virtual repeat rate was not sufficient and that he would type faster with individual physical key-presses. Given that participants 5 and 7 showed such different behavior and the small number of participants overall, the use of mixed effects models for the analysis would not have been effective, requiring a very complicated modeling exercise. As a consequence, evaluation focused only on the differences in TE rates at the start (Session 1), midpoint (Session 10), and end (Session 20) of the longitudinal study.

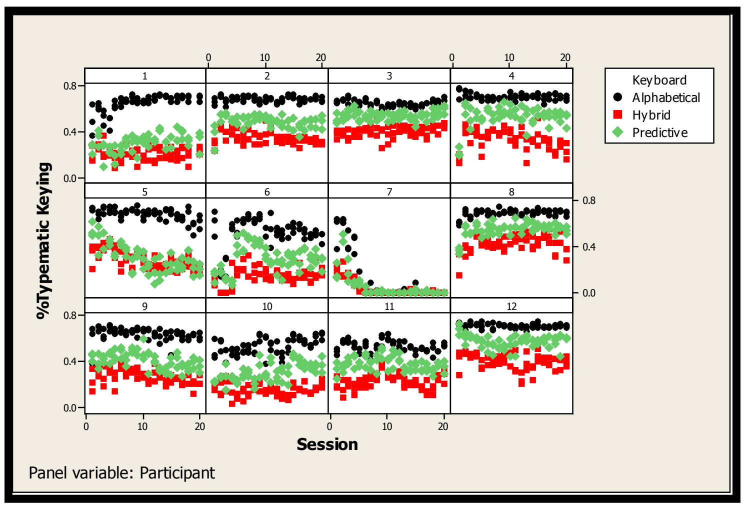


Figure 6-54 Typematic Events for each Individual across Sessions

The mean typematic keying (TE) rates across keyboards, for Session 1, were 60.9% for Alphabetical, 35.3% for Predictive, and 24.6% for the Hybrid keyboard. Table 6-86 provides descriptive statistics for TE by Keyboard. A Friedman test indicated that the rate of typematic keying for Session 1 was different across keyboards, $X^2(2)= 22.167$, $p< .001$. A post-hoc analysis based on Friedman rank-averages indicated significantly greater typematic keying rates with the Alphabetical over the Hybrid ($p< .001$) keyboard. The remaining rank averages were not significantly different.

Variable	Keyboard	N	M	SD	Min	Q1	Mdn	Q3	Max
TE	Alphabetical	36	0.609	0.117	0.360	0.494	0.639	0.701	0.777
	Predictive	36	0.353	0.166	0.089	0.215	0.336	0.469	0.726
	Hybrid	36	0.246	0.105	0.063	0.158	0.220	0.334	0.477

Table 6-86 Descriptive Statistics for Typematic Events by Keyboard for Session 1

At Session 10, the mean typematic keying (TE) rates across keyboards were 58.4% for Alphabetical, 39.6% for Predictive, and 27.7% for the Hybrid keyboard. Table 6-87 provides descriptive statistics for TE by Keyboard. A Friedman test indicated that rate of typematic keying at Session 10 was different across keyboards, $X^2(2)= 18.50$, $p< .001$. As for Session 1, post-hoc analysis, based on Friedman rank-averages, indicated significantly greater typematic keying rates with the Alphabetical over the Hybrid ($p< .0001$) keyboard. The remaining rank averages were not significantly different.

Variable	Keyboard	N	M	SD	Min	Q1	Mdn	Q3	Max
TE	Alphabetical	36	0.584	0.199	0	0.599	0.656	0.696	0.737
	Predictive	36	0.396	0.166	0	0.292	0.422	0.531	0.637
	Hybrid	36	0.277	0.139	0	0.172	0.287	0.379	0.497

Table 6-87 Descriptive Statistics for Typematic Events by Keyboard for Session 10

In Session 20, the mean typematic keying rates across keyboards were 58.3% for Alphabetical, 38.2% for Predictive, and 25.5% for Hybrid. Table 6-88 provides descriptive statistics for TE by Keyboard. A Friedman test indicated that the rate of typematic keying was different across keyboards, $X^2(2)= 18.77$, $p <.0001$. As for Session 1 and 10, post-hoc analysis, based on Friedman rank-averages, indicated significantly greater typematic keying rates of use with the Alphabetical over the Hybrid ($p< .0001$) keyboard. The remaining rank averages were not significantly different.

Variable	*Keyboard*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
TE	Alphabetical	36	0.583	0.191	0	0.560	0.655	0.696	0.716
	Predictive	36	0.382	0.177	0	0.296	0.392	0.540	0.622
	Hybrid	36	0.255	0.118	0	0.200	0.243	0.323	0.476

Table 6-88 Descriptive Statistics for Typematic Events by Keyboard for Session 20

Typematic keying, across all keyboards, were expected to increase between Session 1 and 10, and also between Session 10 and 20. With learning, participants were expected to become familiar with the layouts and to employ typematic keying more often. However, Friedman tests showed no significant effect between sessions for Hybrid ($X^2(2)= 2.09$, $p= .353$), Predictive ($X^2(2)= 3.50$, $p= .174$), or Alphabetical ($X^2(2)= 0.30$, $p= .862$) keyboards, indicating that any possible learning effects did not intensify the use of typematic keying throughout the experiment.

6.4.5 Preference

6.4.5.1 Keyboard Ranking

The mean ranks for keyboard layout, at the last session, were 2.75 for the Alphabetical keyboard, 1.92 for the Predictive keyboard, and 1.33 for the Hybrid keyboard (a lower mean rank is better). Figures 6-55 and 6-56 show the mean ranking results across sessions. Participants' comments describing the reasons for their rankings, across sessions, appear in Appendix H.

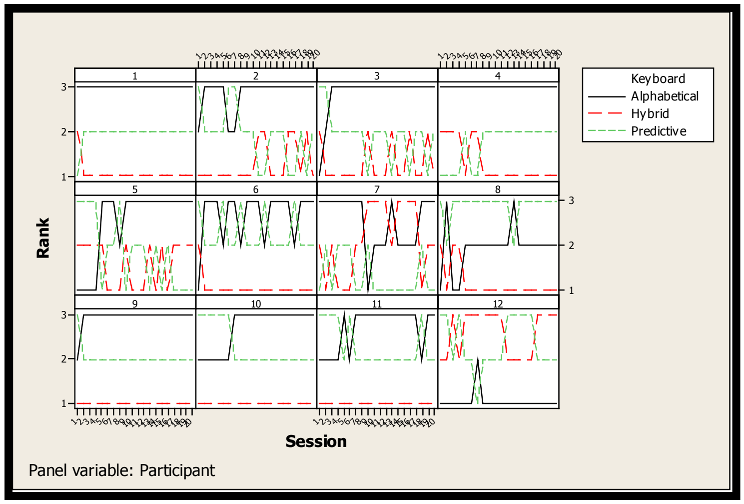


Figure 6-55 Keyboard Rank across Sessions by Participant

As is evident in Figure 6-56, the Hybrid keyboard, on average, was the most preferred keyboard across all sessions. Analysis with a Friedman test showed a significant effect of keyboard across sessions (($X^2(2)= 37.70$, $p <.001$). A post-hoc analysis based on Friedman rank-averages indicated that participants significantly preferred the Hybrid keyboard over the Predictive keyboard ($p <.005$) and the Alphabetical keyboard ($p <.001$), and preferred the Predictive keyboard over the Alphabetical keyboard ($p <.025$). A Friedman test also showed no significant effect of session for the Hybrid keyboard ($X^2(19)= 19.48$, $p =.426$). This indicated that learning effects did not affect preference for the Hybrid layout throughout the experiment.

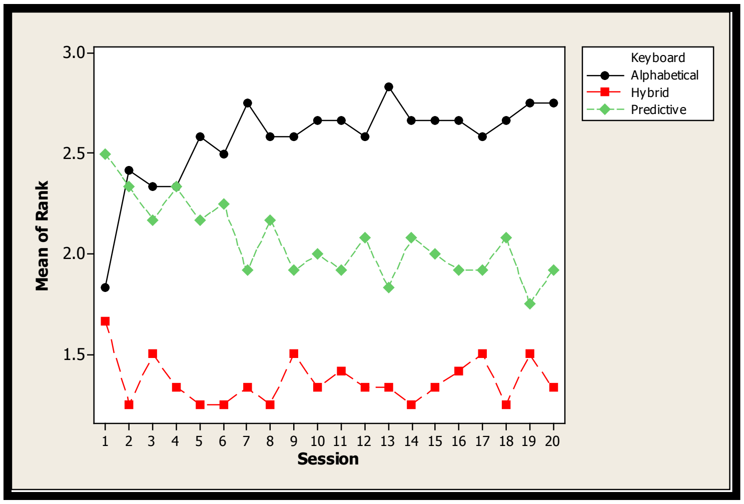


Figure 6-56 Mean Keyboard Rank across Sessions

Friedman tests showed a significant effect of session for the Predictive ($X^2(19)$= 30.90, p =.041) and Alphabetical ($X^2(19)$= 45.76, p <.001) keyboards. Multiple Wilcoxon Signed-Rank tests were used to follow up this finding. A multiple testing correction was applied (α= .01). Preference for the Predictive keyboard did not differ at the onset of learning, compared to the median rank for the Alphabetical keyboard over the same period (T= 25.50, p = .272). There was a suggestive preference for the Predictive over the Alphabetical keyboard at Sessions 10 and 20 (see Table 6-89).

Session	***Wilcoxon Signed-Rank Test Results***
1	T= 25.50, p = .272
10	T= 13.00, p = .021
20	T= 11.00, p = .019

Table 6-89 Test for Preference between Predictive and Alphabetical Keyboards across Sessions 1, 10, and 20

6.4.5.2 Keyboard Rating Data

The rating for each layout was the mean of the six items on the Keyboard Layout Rating Form (see section 5.5.4), which participants completed after typing the text phrases with each keyboard at the end of each session. The questionnaire used 7 point scales, with lower ratings better than higher ratings. The ratings provide an additional indicator of preference.

The mean ratings across all 20 sessions were 2.715 (SD= 1.13) for Alphabetical, 3.365 (SD= 1.10) for Predictive, and 2.611 (SD= 1.09) for Hybrid. At the last session, mean keyboard ratings were 2.458 (SD= 1.05) for Alphabetical, 2.958 (SD= 1.15) for Predictive, and 2.333 (SD= 1.10) for Hybrid. Tables 6-90 and 6-91 provide descriptive statistics for ratings by keyboard layout across all sessions and at Session 20, respectively. Figure 6-57 depicts the mean keyboard ratings by session and keyboard layout.

Variable	*Keyboard*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
$Rating_{All}$	Alphabetical	240	2.715	1.130	1.000	1.833	2.667	3.333	6.167
	Predictive	240	3.365	1.098	1.167	2.500	3.167	4.000	6.333
	Hybrid	240	2.611	1.086	1.000	1.833	2.500	3.333	5.833

Table 6-90 Descriptive Statistics for Keyboard Ratings across all Sessions

Variable	*Keyboard*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
$Rating_{20}$	Alphabetical	12	2.458	1.054	1.167	1.375	2.500	3.167	4.333
	Predictive	12	2.958	1.148	1.333	2.042	3.000	3.625	5.000
	Hybrid	12	2.333	1.105	1.000	1.334	2.167	2.958	4.833

Table 6-91 Descriptive Statistics for Keyboard Ratings at Session 20

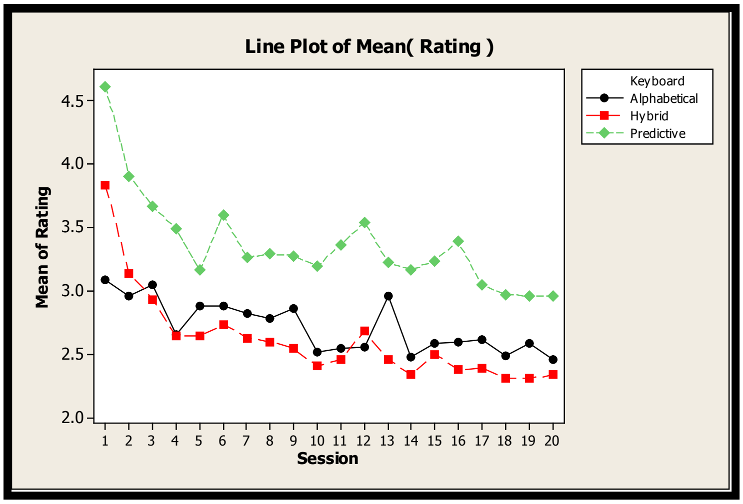


Figure 6-57 Keyboard Rating for the Alphabetical, Predictive, and Hybrid Keyboards across Sessions

It was hypothesized that subjective ratings for the Hybrid keyboard would exceed those for the Alphabetical and Predictive keyboards after only 90 minutes of practice (~ at Session 10). Pairwise t-tests, however, indicated that there was no significant difference in rating between the Hybrid and Alphabetical keyboards ($t(11) = -.226$, $p = .825$). The Predictive keyboard had significantly worse ratings at Session 10 than the Alphabetical ($t(11) = -2.380$, $p = .037$) and Hybrid ($t(11) = 2.318$, $p = .041$) keyboards.

A mixed effects model was employed for the analysis of the longitudinal data. As depicted in Figure 6-58, there was a non-linear relationship between keyboard rating and session. A log transformation minimized the non-linearity in the profiles and normalized the residuals. The best fitting model to the data, among the models considered, was a quadratic model. This model was selected using a forward selection procedure and by inspection of associated residual plots and goodness-of-fit criteria.

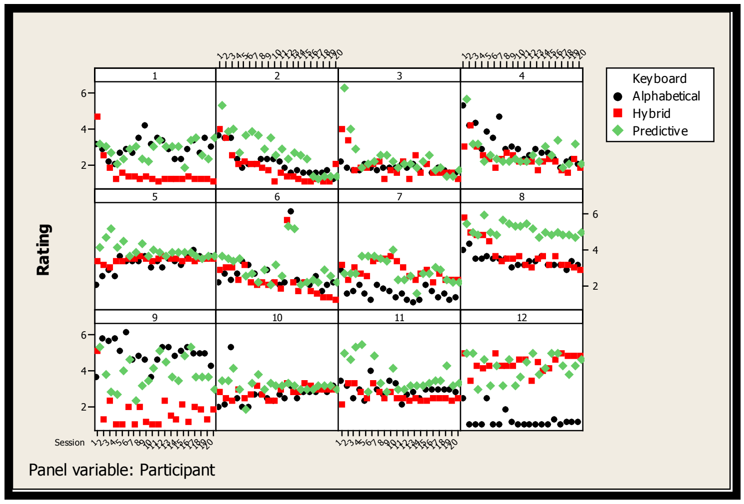


Figure 6-58 Individual Value Plot of Keyboard Rating across Sessions for each Participant

The following mixed effects quadratic model was selected for changes in rating:

(1) $\text{Log(Rating)} = 1.092360 + 0.302923*\text{Keyboard}_P + 0.053054*\text{Keyboard}_H - 0.029663*\text{Session} - 0.005384*\text{Keyboard}_P*\text{Session} - 0.009172*\text{Keyboard}_H*\text{Session} + 0.000906*\text{Session}^2$

Therefore, the estimated average equations for changes in ratings for each keyboard were:

(2) $\text{Log(Rating)}_{\text{Alphabetical}} = 1.092360 - 0.029663*\text{Session} + 0.000906*\text{Session}^2$

(3) $\text{Log(Rating)}_{\text{Predictive}} = 1.395283 - 0.035047*\text{Session} + 0.000906*\text{Session}^2$

(4) $\text{Log(Rating)}_{\text{Hybrid}} = 1.145414 - 0.038835*\text{Session} + 0.000906*\text{Session}^2$

Correlation within participant was captured via an auto regressive model with lag 1 (AR1). Indicator variables were chosen for keyboard as the learning curves obtained are specific to the levels selected. Keyboard had three levels modeled by using two indicator variables, Keyboard_P and Keyboard_H, for the Predictive and Hybrid keyboards, respectively. The coefficients for these variables were compared to the baseline Alphabetical keyboard. Age, gender, and experience did not have a significant effect on changes in keyboard preference.

The response versus fitted value plot appears in Figure 6-59. The plot shows an adequate model fit to the ratings data. The standard deviations for the random effect terms are in Table 6-92. Collectively, the standard deviations of the intercept, session, and keyboard random effects explained a large portion of the variability in the observed data when compared to the standard deviation of the error residual term. There were high negative correlations between the intercept and Predictive keyboard effects (-0.871) and Hybrid keyboard effects (-0.888), but a low correlation between session and Predictive keyboard effects (0.138) and Hybrid keyboard effects (0.116). There was a high correlation (0.948) between the Predictive and Hybrid keyboard effects.

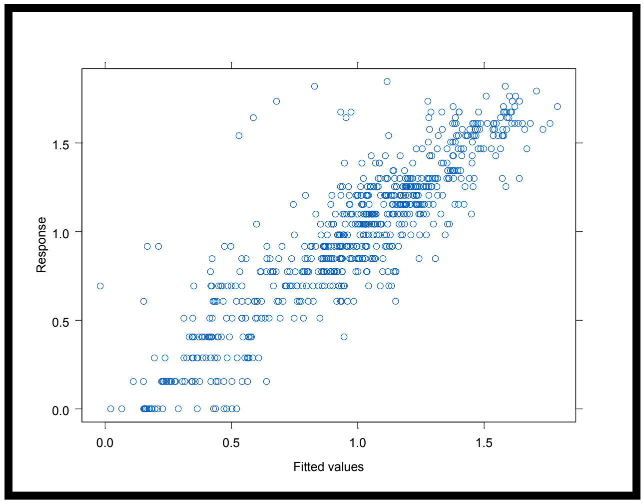


Figure 6-59 Response vs. Fitted Value Plot of the Model for Rating

Parameter	***Std Dev***	***Correlation Matrix***		
		Intercept	**Session**	**$Keyboard_P$**
Intercept	0.382517			
Session	0.016986	-0.200		
$Keyboard_P$	0.389976	-0.871	0.138	
$Keyboard_H$	0.611941	-0.888	0.116	0.948
Residual	0.205917			

Table 6-92 Estimates of the Variance Components of the Random Effect Terms of the Model for Rating

Table 6-93 shows the model parameter estimates and the associated *p*-values. The average initial ratings for the keyboards studied were 1.09 for the Alphabetical, 1.40 for the Predictive, and 1.14 for the Hybrid. Among the main effects, the Predictive keyboard was rated worse than the Alphabetical ($t(702)= 2.551$, $p= .011$) and Hybrid keyboards ($F(1,702)= 8.2324$, $p= .004$) at the onset of learning.

Parameter	*Value*	*Std. Error*	*DF*	*t-value*	*p-value*
Intercept	1.092360	0.155735	702	9.438474	<.0001
Session	-0.029663	0.007907	702	-3.751533	0.0002
$Session^2$	0.000906	0.000275	702	3.294509	0.0010
$Keyboard_P$	0.302923	0.118763	702	2.550656	0.0110
$Keyboard_H$	0.053054	0.180625	702	0.293726	0.7691
$Keyboard_P$* Session	-0.005384	0.003160	702	-1.703612	0.0889
$Keyboard_H$*Session	-0.009172	0.003145	702	-2.916095	0.0037

Table 6-93 Parameter Estimates for the Ratings Quadratic Model using Session

The data indicated a significant difference in preference ratings over time between the Hybrid and Alphabetical keyboards (t (702) =-2.916, p=.004). This implies that over time, ratings for the Hybrid keyboard improved more quickly than those for the Alphabetical keyboard. The model results did not indicate any difference in changes for preference ratings between the Predictive and Hybrid keyboards (F(1,702)=1.450, p=.229). This means that initial average ratings when typing with the Predictive keyboard were worse, but attitudes about the keyboards improved at similar rates for the Predictive and Hybrid keyboards. Figure 6-60 presents the estimated profiles based on the fitted model. Preference ratings in time tended to decrease (i.e. improve) over time for most participants. For some, the decrease was more rapid, for others, less so.

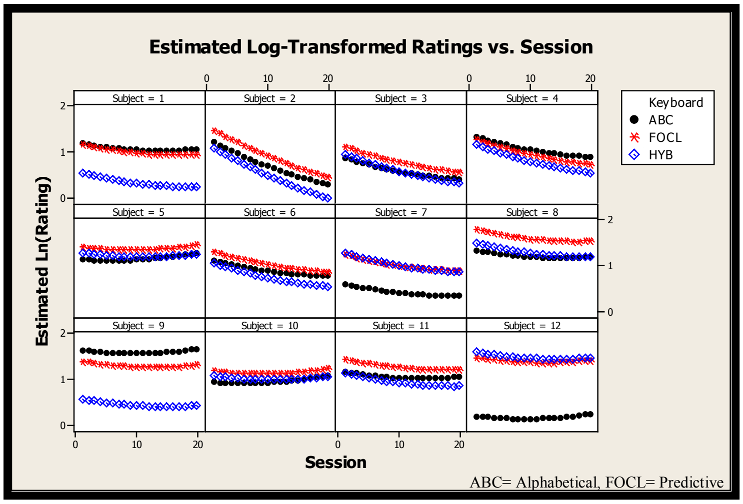


Figure 6-60 Estimated Learning Trends for Rating

6.4.5.3 Keyboard Rank vs. Rating

As in the previous experiment, there appeared to be a lack of correspondence between participant rank and ratings for the keyboards. The ranking data indicated a clear preference for the Hybrid keyboard over Alphabetical and Predictive. However, this was not fully captured by the keyboard ratings. Pearson correlation coefficients were calculated for all pairs of keyboard rank and ratings (Alphabetical Rank/Rating ($r = .294$, $p < .001$), Hybrid Rank/Rating ($r = .473$, $p < .001$), and Predictive Rank/Rating ($r = .544$, $p < .001$). Surprisingly, there were highly significant relationships between the rank and rating scores for all keyboards. These results provide additional evidence to support the further investigation of this discrepancy to ascertain opportunities for improving the rating questionnaire.

6.4.6 Keyboard Attribute Importance

The rating for keyboard attributes was based on six items on the Keyboard Attribute Importance Rating Form (see section 5.5.4), given at the start of each session. The questionnaire used 7 point scales, with higher ratings indicative of greater importance. Overall, participants rated all questionnaire attributes as important, with average importance ratings exceeding 5.5 for all attributes.

The mean attribute importance ratings for Session 1 were 5.67 for acceptability of keyboard layout, 6.25 for ease of finding letters, 5.75 for ease of learning letter locations, 5.92 for ease of typing, 5.92 for ease of typing accurately, and 5.50 for ease of typing letters. Analysis with a Friedman test indicated that importance rating for Session 1 was not different across attributes, $X^2(5)= 5.25$, $p = .386$. Table 6-94 provides descriptive statistics for importance rating by attribute.

Variable	*Keyboard*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
Importance Rating	Acceptability of keyboard layout	12	5.667	0.985	4	5.00	6.00	6.00	7
	Ease of finding letters	12	6.250	0.622	5	6.00	6.00	7.00	7
	Ease of learning letter locations	12	5.750	1.138	4	5.00	6.00	7.00	7
	Ease of typing	12	5.917	0.996	4	5.00	6.00	7.00	7
	Ease of typing accurately	12	5.917	0.900	4	5.25	6.00	6.75	7
	Ease of typing letters	12	5.500	1.000	4	5.00	5.50	6.00	7

Table 6-94 Descriptive Statistics for Importance Rating by Attribute at Session 1

At Session 10, the mean importance rating across attributes was 5.83 for acceptability of keyboard layout, 6.25 for ease of finding letters, 5.92 for ease of learning letter locations, 5.92 for ease of typing, 6.00 for ease of typing accurately, and 6.08 for ease of typing letters. A Friedman test indicated that importance rating at Session 10 was not statistically different across attributes ($X^2(5)= 4.74$, $p =.448$). Table 6-95 provides descriptive statistics for importance rating by attribute.

Variable	*Keyboard*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
Importance Rating	Acceptability of keyboard layout	12	5.833	0.835	4	5.25	6.00	6.00	7
	Ease of finding letters	12	6.250	0.754	5	6.00	6.00	7.00	7
	Ease of learning letter locations	12	5.917	0.793	4	6.00	6.00	6.00	7
	Ease of typing	12	5.917	0.669	5	5.25	6.00	6.00	7
	Ease of typing accurately	12	6.000	0.953	4	5.25	6.00	7.00	7
	Ease of typing letters	12	6.083	0.900	4	6.00	6.00	7.00	7

Table 6-95 Descriptive Statistics for Importance Rating by Attribute at Session 10

For Session 20, the mean importance ratings across attributes were 5.92 for acceptability of keyboard layout, 6.25 for ease of finding letters, 5.92 for ease of learning letter locations, 6.00 for ease of typing, 6.25 for ease of typing accurately, and 5.92 for ease of typing letters. A Friedman test indicated that importance rating at Session 20 was also not significantly different across attributes ($X^2(5)= 6.95$, $p =.224$). Table 6-96 provides descriptive statistics for importance rating by attribute.

Variable	*Keyboard*	*N*	*M*	*SD*	*Min*	*Q1*	*Mdn*	*Q3*	*Max*
Importance Rating	Acceptability of keyboard layout	12	5.917	1.084	4	5.25	6.00	7.00	7
	Ease of finding letters	12	6.250	0.754	5	6.00	6.00	7.00	7
	Ease of learning letter locations	12	5.917	0.793	4	6.00	6.00	6.00	7
	Ease of typing	12	6.000	1.128	4	5.25	6.00	7.00	7
	Ease of typing accurately	12	6.250	0.866	5	5.25	6.00	7.00	7
	Ease of typing letters	12	5.917	0.996	4	6.00	6.00	6.00	7

Table 6-96 Descriptive Statistics for Importance Rating by Attribute at Session 20

Furthermore, Friedman tests showed no significant effect between Sessions (1,10,20), in importance rating for acceptability of keyboard layout ($X^2(2)= 0.00$, $p= 1.000$), for ease of finding letters ($X^2(2)= 5,43$, $p= .066$), for ease of learning letter locations ($X^2(2)= 1.75$, p= .418), for ease of typing ($X^2(2)$= 0.35, p= .840), for ease of typing accurately ($X^2(2)$= 0.08, p= .961), or for ease of typing letters ($X^2(2)$= 0.86, p= .651).

At the end of the study, participants ranked the attributes in order of importance. However, an analysis of ranks was not significant (Friedman test, X^2 (5) = 9.95, p= .077, see Figure 6-61).

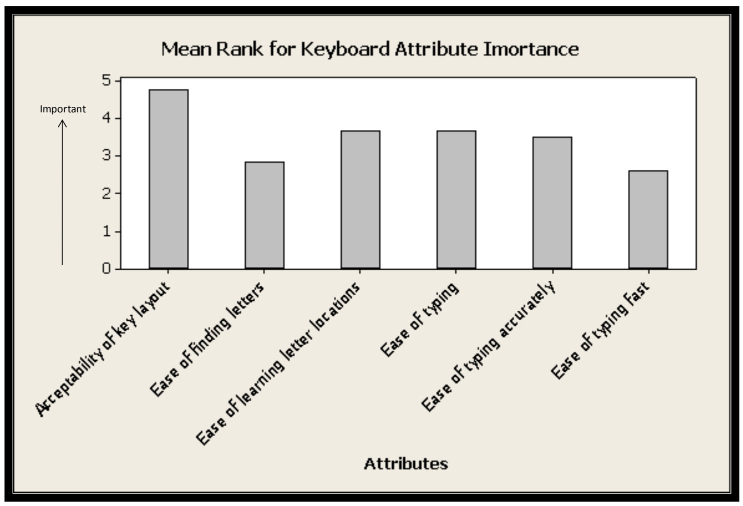


Figure 6-61 Mean Rank for Keyboard Importance Attributes

As in prior studies (e.g. Lewis 1995), the participants considered all of the attributes important. Given the lack of correspondence between keyboard rank and ratings, future work should investigate what additional attributes might lead to greater consistency between keyboard ranking and ratings.

6.4.7 Discussion

This experiment addressed the learning process in typing with the five-key methods. Mean text entry throughputs, across sessions, were 6.87, 7.73, and 7.79 (in CWPM) for the Alphabetical, Predictive, and Hybrid keyboards, respectively. Maximum session averages were 10.612 for Alphabetical at session 13, 13.169 at session 13 for Predictive, and 14.039 at session 20 for Hybrid.

Mixed effect modeling was used to analyze the longitudinal data. Typically, power models are used in modeling learning (Heathcote, Brown and Mewhort, 2009). In fact, most text entry researchers have assumed a power function. Interestingly, the performance data in this study were fitted best with a quadratic function.

Overall, text entry rates increased over time for all participants. For some, the increase was more rapid than for others. Individual trajectories suggest that greater text entry rates for the keyboards could be attained with time, as leveling off was not observed. However, claims for identification of an upper limit are not possible within this experiment given the small sample size and that an asymptote was not apparent in the 20 sessions that made up the follow-up time for this experiment.

Results suggest that the nonstandard keyboards start higher and appear to have higher upper limits than the Alphabetical keyboard. Furthermore, learning rates are also greater for the nonstandard keyboards than the Alphabetical. Initial speeds and learning rates do not differ between the nonstandard keyboards.

The Hybrid keyboard was expected to outperform the Predictive keyboard, given its low KSPC and anticipated reduction in visual search demands. However, post-study interviews revealed that deciding on the navigation path to the intended letter required more effort when typing with the Hybrid keyboard over the Predictive and Alphabetical keyboards. This, along with the results of the reaction time analyses, suggests that this unexpected cognitive demand impeded Hybrid text entry performance.

In general, reaction time dropped across all keyboards as a function of practice. However, there was no evidence that this decrease was associated with any one keyboard. Therefore, the reduction in KSPC for the Hybrid had a corresponding cognitive demand (initial path selection) that eliminated its expected gains over the Predictive. The net effect is that both nonstandard keyboards performed equally, and both better than the Alphabetical keyboard.

Keyboard designs were optimized using KSPC, with a goal for keystroke reduction. In this experiment, the average initial KSPC for the Hybrid was lower than the initial KSPC for the Predictive and Alphabetical keyboards. However, KSPC learning rates were greatest for the Alphabetical keyboard.

The findings for KSPC learning are echoed in the selector movement efficiency results. Specifically, participants achieved greatest relative efficiency with the Alphabetical keyboard, mainly because selecting the optimal path is most apparent with this method. To optimally access target letters from n through z, participants moved the cursor across the onscreen keyboard using the left key, employing the cursor wraparound feature. To optimally access target letters between a and m, participants moved the cursor using the right key. Selecting the optimal path was not as evident for the other methods.

Typematic keying was used extensively by most participants in this study. However, participant 7 stopped using typematic keying altogether after only a few sessions. The participant indicated that the rate for the auto repeats was too slow and that he expected to achieve greater speeds using only physical key presses. Further research should establish the optimal delay time and repeat intervals for selection based text entry systems using typematic keying.

Typematic keying was used to a greater extent with the Alphabetical than with the nonstandard keyboards. This is expected given the larger KSPC and cursor distances inherent to these layouts. However, contrary to expectations, learning effects did not intensify the use of typematic keying with any keyboard.

On average, the Hybrid keyboard was the most preferred keyboard across all sessions. Unexpectedly, results indicated that participants preferred the Predictive keyboard over the Alphabetical. Even though the Alphabetical layout had the greatest relative efficiency, participants' comments indicated that the Alphabetical keyboard was the most frustrating to use due to its inherently higher KSPC.

The average initial ratings for the keyboards studied were 1.09 for the Alphabetical, 1.40 for the Predictive, and 1.14 for the Hybrid, indicating that the Predictive keyboard was rated worse than the Alphabetical and Hybrid keyboards at the onset of learning. Trend in time data suggested that the ratings for the Hybrid keyboard improved more quickly than those for the Alphabetical keyboard, but not for the Predictive.

Chapter 7. Conclusion

Selection based methods, such as five-key techniques, are prime candidates for limited text entry on nonstandard and input-constrained devices. These techniques are widely used on consumer products, such in-car navigation systems, television remotes and gaming controllers. Therefore, the main goal of this dissertation was to create and validate novel and effective five-key text entry techniques for constrained devices.

This research examined the design of alternative keyboard layouts used for five-key text entry techniques. Three keyboard layouts (Alphabetical, Predictive, and Hybrid) were selected to represent standard and less familiar arrangements. The nonstandard layouts are dynamic, using digram prediction to reduce keystrokes. The analysis centered on a series of controlled experiments; entering four different types of texts at three different levels of user training, conducted on a research platform developed by the author.

The immediate usability of three alternative keyboard layouts for supporting five-key text entry was investigated. Results indicated there were no statistically significant differences in performance across the three tested keyboards. Furthermore, experimental results show that following immediate usability, but still at the onset of learning, there was no overall difference in performance among the three keyboard layouts across four text types. However, the Alphabetical keyboard surpassed both the Predictive and Hybrid keyboards in text entry speed in typing Web addresses. The nonstandard keyboards performed superior to the Alphabetical keyboards in typing Words/Spaces and Sentences, but performed no better in typing Address strings than the Alphabetical.

Longitudinal data indicated that the initial speeds and learning rates for the nonstandard (or predictive) layouts exceeded those of the standard (Alphabetical) layout. This is contrary to the existing literature, which generally indicates that learning of nonstandard layouts takes longer than a standard layout (as in Butterbaugh & Rockwell, 1982).

It was expected that the Hybrid keyboard would be superior to the Predictive keyboard, given that the Hybrid keyboard had the greatest reduction in keystrokes and, due to its limited number of predicted next characters, would have reduced demand on visual search. However, overall speeds and error rates did not differ between the nonstandard keyboards. This was due to an unexpected cognitive cost incurred by the Hybrid keyboard. Even though the design efforts for the Hybrid focused on reduction in visual search demands, the proposed layout resulted in inherently greater planning for execution activities as participants decided which initial Hybrid path to pursue- the predicted set or the alphabetic set. Even with the unexpected cost, participants consistently ranked the Hybrid keyboard as the most preferred layout.

This dissertation used hierarchical linear modeling, specifically mixed effects modeling, to assess learning rates. Mixed effects modeling has been used widely in medical and psychology studies, but has not been used in the text entry research.

Using mixed effects modeling provides a significant contribution in evaluating the effects of keyboard learning over time. Mixed effects modeling is a potent statistical method that can be applied to longitudinal research to evaluate an intervention at multiple levels. Typically, repeated measures ANOVAs are used in the text entry literature to evaluate longitudinal data. However, the use of repeated measures ANOVA may have not been appropriate. The major differences between repeated-measures ANOVA and mixed effects modeling are that mixed effects modeling (a) has fewer strict assumptions, (b) has more flexible data requirements (dealing with missing data and collection points that are unequal), and (c) emphasizes individual change over group differences. Given the use of mixed effects modeling, the models presented herein have descriptive ability and some predictive use in comparing the text entry methods.

Furthermore, the use of the power law to model learning is ubiquitous in the text entry research. Most researchers assume the power function rather than testing to determine if it provides a better description than other functions. However, in this research the quadratic model provided the best fit for the skill acquisition data. However, none of the models reached the expected asymptotic expert rates. The quadratic model implicitly suggests that the increasing trends observed for text entry performance will eventually degrade, but the current data are insufficient to support this inherent claim. This means the quadratic model is probably better for describing the experimental results, but may have limited reliability as a predictor.

7.1 Contributions

This dissertation makes the following contributions:

1. Comprehensive review and organization of recent literature on mobile text entry techniques with a focus on the application of selection based methods for constrained devices.
2. Identification of the design factors and interactions that influence the design of selection-based keyboards.
3. Design and development of three keyboard layouts for five-key text entry techniques.

4. Demonstration of the benefits of KSPC as a tool for a priori analyses of text entry techniques. Using KSPC, proposed text entry methods can undergo analysis, comparison, and redesign prior to implementation and evaluation.
5. Development of an extensible and interactive research platform designed to reduce time and effort in the development of prototypes of alternative selection based text entry schemes and their empirical evaluation.
6. Empirical investigation of the usability of the developed keyboard layouts (Alphabetic, Predictive, and Hybrid) at three different levels of user training.
7. Use of mixed effects modeling to perform longitudinal analyses leading to a richer understanding of the learning process and capturing changes in time.

7.2 Future Work

This section details related future work.

7.2.1 Use of Keystroke Level Modeling

Use of KLM provides insight into potential expert user behavior with alternative design configurations. It may be beneficial to use predictive modeling techniques, such as Keystroke Level Modeling (KLM), to improve the Hybrid layout.

7.2.2 Employing Trigram Prediction

In this work, the prediction layouts relied on digram probabilities. It may beneficial to explore performance and preference gains in using trigram or word based disambiguation.

7.2.3 Focusing on Cognitive and Perceptual Demands

Most text entry research has relied on optimization of movement based on the assumption that there is no visual search time for expert users. As is evident in the results of this work, cognitive demands can prevent techniques from reaching maximum performance. Therefore, subsequent research efforts should include assessing the visual attention and cognitive demands for selection based methods.

7.2.4 Improving the Keyboard Rating Questionnaire

There appears to be a lack of correspondence between participant keyboard rank and ratings for the keyboards. This discrepancy should be further investigated to ascertain opportunities for improving the rating questionnaire.

7.2.5 Investigating Typematic Keying Rates

Typematic keying is used extensively with selection based method. However, some participants indicated that the repeat rate was insufficient. Further research is warranted to establish the optimal delay time and repeat intervals for selection based text entry systems using typematic keying.

References

Bellman, T., & MacKenzie, I. S. (1998). A probabilistic character layout strategy for mobile text entry. In *Proceedings of Graphics Interface '98* (pp. 168-176). Vancouver, Canada: Canadian Information Processing Society.

Butterbaugh, L.C., & Rockwell, T.H. (1982). Evaluation of alternative alphanumeric keying logics. *Human Factors,* 24, 521-533.

Butts, L., & Cockburn, A. (2001). An evaluation of mobile phone text input methods. In Grundy, J. & Calder, P. (Eds.), Conferences in Research and Practice in Information Technology: Vol. 7. *Third Australasian User Interfaces Conference* (pp. 55-59). Melbourne, Australia: Australian Computer Society, Inc.

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

Card, S. K., Moran, T. P., & Newell, A. (1980). The keystroke-level model for user performance time with interactive systems. *Communications of the ACM* , 23(7), 396-410.

Card, S. K., Moran, T. P., & Newell, A. (1983). *The psychology of human-computer interaction*. Hillsdale, NJ: Lawrence Erlbaum.

Clarkson, E., Clawson, J., Lyons, K., & Starner, T. (2005) An empirical study of typing rates on mini-QWERTY keyboards. In *CHI '05 Extended Abstracts* (pp. 1288-1291). Portland, OR: Association for Computing Machinery.

Cockburn, A., & Siresena, A. (2003). Evaluating mobile text entry with the Fastap keypad. In *People and Computers XVII: Proceedings of the 17th Annual British Computer Society Conference on Human-Computer Interaction* (pp. 77-80). Bath, UK: British Computer SocietyCurran, K., Woods, D., & Riordan, B. O. (2006). Investigating text input methods for mobile phones. *Telematics and Informatics*, *23*, 1-21.

De Jong, J.R. (1957). The effect of increased skills on cycle time and its consequences for time standards. *Ergonomics*, 1, 51–61.

Detweiler, M. C., Schumacher, R. M., Jr., & Gattuso, N. L., Jr. (1990). Alphabetic input on a telephone keypad. In *Proceedings of the Human Factors Society 34th Annual Meeting* (pp. 212-216). Santa Monica, CA: Human Factors Society.

Dunlop, M. D., & Crossan, A. (2000). Predictive text entry methods for mobile phones. *Personal Technologies*, *4*, 134-143.

Dunlop, M. D., & Masters, M. M. (2007). Investigating five key predictive text entry with combined distance and keystroke modeling. *Personal and Ubiquitous Computing*. Retrieved November 25, 2008, from http://www.springerlink.com/content/pq30lk2226408kt4/

Evreinova, T., Evreinov, G. & Raisamo, R. (2004) Four key text entry for physically challenged people. *Adjunct Proc. UI4ALL '04.* http://ui4all.ics.forth.gr/workshop2004/

Fitts, P. M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47(6), 281–391.

Gong, J., & Tarasewich, P. (2005). Alphabetically constrained keypad designs for text entry on mobile phones. *Proceedings of the ACM Conference on Human Factors in Computing Systems - CHI 2005* (pp. 211-220). New York: Association for Computing Machinery.

Hart, S. & Staveland, L. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research, *in* P. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139-183). Amsterdam, Netherlands: Elsevier Science.

Heathcote, A., Brown, S., & Mewhort, D. J. K. (2000). The power law repealed: The case for an exponential law of practice. *Psychonomic Bulletin & Review*, 7(2), 185-207.

Isokoski, P. (1999). A minimal device-independent text input method. Unpublished master's thesis, University of Tempere, Finland.

Isokoski, P. (2001). Model for Unistroke writing time. In *Proceedings of CHI 2001* (pp. 357-364). Seattle, WA: Association for Computing Machinery.

Isokoski, P. (2004). Manual text input: Experiments, models, and systems. Unpublished doctoral dissertation, University of Tempere, Finland.

Isokoski, P., & MacKenzie, I. S. (2003). Combined model for text entry rate development. In *Proceedings of CHI 2003* (pp. 752-753). Ft. Lauderdale, FL: Association for Computing Machinery.

Isokoski, P., & Raisamo, R. (2004). Quikwriting as a multi-device text entry method. *Proceedings of the Third Nordic Conference on Human-Computer Interaction NordiCHI 2004* (pp.105–108). New York: ACM Press.

James, C. L., & Reischel, K. M. (2001). Text input for mobile devices: comparing model prediction to actual performance. In *Proceedings of CHI 2001* (pp. 365-371). Seattle, WA: Association for Computing Machinery.

John, B. E. (2003). Information processing and skilled behavior. In J.M. Carroll (Ed.) HCI models, theories, and frameworks: Toward a multidisciplinary science (pp. 55-102). San Francisco, CA: Morgan Kaufmann.

Koester, H.H., & Levine. S.P. (1994). Modeling the speed of text entry with a word prediction interface. *IEEE Transactions on Rehabilitation Engineering*, 2(3), 177 -187.

Koltringer, T., & Grechenig, T. (2004). Comparing the immediate usability of Graffiti 2 and virtual keyboard. In *Proceedings of CHI 2004 Extended Abstracts* (pp. 1175-1178). New York, NY: Association for Computing Machinery.

Lewis, J.R. (1992). *Typing-key layouts for single-finger or stylus input: Initial user preference and performance* (Technical Report No. 54729). Boca Raton, FL: International Business Machines Corporation.

Lewis, J. R. (1993a). The *Friedman rank sum test with multiple comparisons: A BASIC program* (Tech. Report 54.733). Boca Raton, FL: International Business Machines Corp.

Lewis, J.R. (1993b). 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. (1995). Input rates and user preference for three small-screen input methods: standard keyboard, predictive keyboard, and handwriting (Tech. Report 54.889). Boca Raton, FL: IBM Corp.

Lewis, J. R., Commarford, P. M., Kennedy, P. J., & Sadowski, W. J. (2008). Handheld electronic devices. In C. Melody Carswell (Ed.), Reviews of Human Factors and Ergonomics, Vol. 4 (pp. 105-148). Santa Monica, CA: Human Factors and Ergonomics Society.

Lewis, J. R., Kennedy, P. J., & LaLomia, M. J. (1999). Development of a digram-based typing key layout for single-finger/stylus input. In *Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting* (pp. 415-419). Santa Monica, CA: Human Factors and Ergonomics Society.

Lewis, J. R., LaLomia, M. J., & Kennedy, P. J. (1999). Evaluation of typing key layouts for stylus input. In *Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting* (pp. 420-424). Santa Monica, CA: Human Factors and Ergonomics Society.

Lewis, J. R., Potosnak, K. M., & Magyar, R. (1997). Keys and keyboards. In M. Helander, T. K. Landauer, & P. V. Prabhu (Eds.), Handbook of human-computer interaction (pp. 1285–1315).Amsterdam: North-Holland.

Lowry, R. (2005). Concepts and applications of inferential statistics [Electronic Version]. Web Site:http://faculty.vassar.edu/lowry/webtext.html.

Lyons, K., Plaisted, D., & Starner, T. (2004). Expert chording text entry on the Twiddler one-handed keyboard. In *Proceedings of the 8th International Symposium on Wearable Computers* (pp. 94-101). Arlington, VA: Institute of Electrical and Electronics Engineers.

Lyons, K., Starner, T., Plaisted, D., Fusia, J., Lyons, A., Drew, A., & Looney, E. (2004). Twiddler typing: one-handed chording text entry for mobile phones. In *Proceedings of CHI 2004* (pp. 671-678). Vienna, Austria: Association for Computing Machinery.

MacKenzie, I. S. (2002a). KSPC (keystrokes per character) as a characteristic of text entry techniques. In the *Proceedings of the Fourth International Symposium on Human-Computer Interaction with Mobile Devices* (pp. 195-210). Heidelberg, Germany: Springer-Verlag.

MacKenzie, I. S. (2002b). Mobile text entry using three keys. In *Proceedings of NordiCHI 2002* (pp. 27-34). Arhus, Denmark: Association for Computing Machinery.

MacKenzie, I. S. (2003). Motor behavior models for human-computer interaction. In J.M. Carroll (Ed.) HCI models, theories, and frameworks: Toward a multidisciplinary science. pp. 27-54. San Francisco: Morgan Kaufmann.

MacKenzie, I. S., Kober, H., Smith, D., Jones, T., & Skepner, E. (2001). Letterwise: Prefix-based disambiguation for mobile text input. In *Proceedings of the 14th Annual ACM Symposium on User Interface Software and Technology UIST 2001* (pp. 111-120). Orlando, FL: Association for Computing Machinery.

MacKenzie, I. S., & Soukoreff, R. W. (2002a). A model of two-thumb text entry. In *Proceedings of Graphics Interface 2002* (pp. 117-124). Calgary, Canada: Canadian Information Processing Society.

MacKenzie, I. S., & Soukoreff, R. W. (2002b). Text entry for mobile computing: Models and methods, theory and practice. *Human-Computer Interaction*, *17*, 147-198.

MacKenzie, I.S., & Soukoreff, R.W. (2003). Phrase sets for evaluating text entry techniques. *Extended Abstracts of the ACM Conference on Human Factors in Computing Systems - CHI 2003*, 754-755.

MacKenzie, I. S., & Zhang, S. X. (1997). The immediate usability of Graffiti. In *Proceedings of Graphics Interface1997* (pp. 129–137). Mississauga, Ontario: Canadian Information Processing Society.

MacKenzie, I. S., & Zhang, S. X. (1999). The design and evaluation of a high-performance soft keyboard. In *Proceedings of CHI 1999* (pp. 25-31). Pittsburgh, PA: Association for Computing Machinery.

MacKenzie, I. S., & Zhang, S. X. (2001). An empirical investigation of the novice experience with soft keyboards. *Behaviour and Information Technology*, *20*, 411-418.

MacKenzie, I. S., Zhang, S. X., & Soukoreff, R. W. (1999). Text entry using soft keyboards. *Behaviour and Information Technology*, *18*, 235-244.

Masui, T. (1998). An efficient text input method for pen-based computers. In *Proceedings of CHI 1998* (pp. 328-335). Los Angeles, CA: Association for Computing Machinery.

Matias, E., MacKenzie, I. S., & Buxton, W. (1996). One-Handed Touch Typing on a QWERTY Keyboard. *Human-Computer Interaction*, 11, 1-27.

Mayzner, M.S., & Tresselt, M.E. (1965). Tables of single-letter and digram frequency counts for various word-length and letter-position combinations, *Psychonomic Monograph Supplements, 1(2)*, 13-32.

McQueen, J.C., MacKenzie, I.S. & Zhang, S.X. (1995). An extended study of numeric entry on pen-based computers. In Proceedings of Graphics Interface '95 (pp. 215-222). Canadian Information Processing Society.

Microsoft Excel (2007) [Computer Software]. Redmond, WA: Microsoft Corporation.

Millet, B., Asfour, S., & Lewis, J.R. (2008). Designing a hybrid layout for a five-key text entry technique. In the Proceedings of the *International Society for Occupational Ergonomics and Safety*, Chicago, IL, June 12-13, 2008, pp. 156-162.

Millet, B., Asfour, S., & Lewis, J.R. (2009). Selection-based virtual keyboard prototypes and data collection application. *Behavior Research Methods*, 41(3), 951-956.

Minitab (15) [Computer Software]. State College, PA: Minitab, Inc.

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

Pavlovych, A., & Stuerzlinger, W. (2004). Model for non-expert text entry speed on 12-button phone keypads, In Proceedings of *Conference on Human Factors in Computing Systems 2004* (pp. 351-358). Vienna, Austria: Association for Computing Machinery.

Sandnes, F. E., Thorkildssen, H. W., Arvei, A., & Buverud, J. O. (2004). Techniques for fast and easy mobile text-entry with three-keys. In *Proceedings of the 37th Hawaii International Conference on System Sciences* (pp. 1-10). Hawaii, HI: Institute of Electrical and Electronics Engineers.

Sears, A., Jacko, J. A., Chu, J., & Moro, F. (2001). The role of visual search in the design of effective soft keyboards. *Behaviour & Information Technology*, *20*, 159-166.

Sears, A., & Zha, Y. (2003). Data entry for mobile devices using soft keyboards: Understanding the effects of keyboard size and user tasks. *International Journal of Human-Computer Interaction*, *16*, 163-184.

Silfverberg, M. (2007). Historical overview of consumer text entry technologies. In I. S. MacKenzie, & K. Tanaka-Ishii (Eds.), *Text entry systems: Mobility, accessibility, universality* (pp. 3-25). San Francisco, CA: Morgan Kaufmann.

Silfverberg, M., MacKenzie, I. S., & Korhonen, P. (2000). Predicting text entry speed on mobile phones. In *Proceedings of CHI 2000* (pp. 9-16). Amsterdam, Netherlands: Association for Computing Machinery.

Sirisena, A. (2002). *Mobile text entry.* Unpublished master's thesis, University of Canterbury , Christchurch, New Zealand.

Soukoreff, R. W., & MacKenzie, I. S. (1995). Theoretical upper and lower bounds on typing speed using a stylus and soft keyboard. *Behaviour & Information Technology*, *14*, 370-379.

Soukoreff, R. W., & MacKenzie, I. S. (2003a). Input-based language modeling in the design of high performance input systems. Proceedings of Graphics Interface 2003 (pp. 89-96). Toronto: Canadian Information Processing Society.

Soukoreff, R. W., & MacKenzie, I.S. (2003b). Metrics for text entry research: An evaluation of MSD and KSPC, and a new unified error metric. *Proceedings of the ACM Conference on Human-Factors in Computing Systems - CHI 2003*, 113-120.

S Plus (8.0) [Computer Software]. Sommerville, MA: TIBCO Software Inc.

SPSS (17.0) [Computer Software]. Chicago, IL: SPSS, Inc.

Ward, D. J., Blackwell, A. F., & MacKay, D. J. C. (2002). Dasher: A gesture-driven data entry interface for mobile computing. *Human-Computer Interaction*, *17*, 199-228.

Wigdor, D., & Balakrishnan, R. (2003). TiltText: Using tilt for text input to mobile phones. In *Proceedings of the 16th Annual ACM Symposium on User Interface Software and Technology* (pp. 81–90).New York: Association for Computing Machinery.

Wobbrock, J. O., Myers, B. A., & Rothrock, B. (2006). Few-key text entry revisited: Mnemonic gestures on four keys. In *Proceedings of the ACM Conference on Human Factors in Computing Systems '06* (pp. 489-492). New York, New York: ACM Press.

Wobbrock, J.O. (2007). Measures of text entry performance. In I. S. MacKenzie, & K. Tanaka-Ishii (Eds.), *Text entry systems: Mobility, accessibility, universality* (pp. 47-74). San Francisco, CA: Morgan Kaufmann.

Zhai, S., Hunter, M., & Smith, B. A. (2000). The Metropolis keyboard – an exploration of quantitative techniques for virtual keyboard design. In *Proceedings of ACM Symposium on User Interface Software and Technology* (pp. 119-128). San Diego, CA: Association for Computing Machinery.

Zhai, S., Hunter, M., & Smith, B. A. (2002). Performance optimization of virtual keyboards. *Human-Computer Interaction*, *17*, 89-129.

Zhai, S., Kristensson, P.-O., & Smith, B. A. (2005). In search of effective text input interfaces for off the desktop computing. *Interacting with Computers*, *17*, 229-250.

Zhai, S., Sue, A., Accot, J. (2002). Movement model, hits distribution and learning in virtual keyboarding. *Proceedings of the CHI'2002*: ACM Conference on Human Factors in Computing Systems, CHI Letters 4 (1), 17-24.

Appendix A. Pre-Test Questionnaire

Thank you for your participation. Before we begin, we would like to get to know a little bit more about you and your current use of handheld technologies. Please take a few minutes to answer this questionnaire.

1) Age:
   ( ) < 25 years   ( ) 25 – 34 years   ( ) 35 – 44 years
   ( ) 45 – 54 years   ( ) 55 – 64 years   ( ) > 65 years

2) Gender:
   ( ) Male   ( ) Female

3) Are you
   ( ) Right Handed   ( ) Left Handed

4) Education Level (select highest level achieved):
   ( ) High School Graduate   ( ) Vocational/ Technical Graduate
   ( ) Some College   ( ) Bachelors Degree
   ( ) Masters Degree   ( ) Doctoral Degree
   ( ) Other: ______________________

5) What is your job title?
   Job title: ____________________________________

6) Do you have a mobile phone? (Work or personal)
   ( ) Yes (Work)
   ( ) Yes (Personal)
   ( ) Both
   ( ) No

7) What features have you used on any mobile device? (Check all that apply)
   ( ) Calling/ receiving
   ( ) Texting
   ( ) Emailing
   ( ) Push to Talk (PTT)
   ( ) Other: ______________________________

8) Have you ever used a PDA (i.e. Palm Pilot) or smart phone?
   ( ) Yes
   ( ) No

9) Your current phone is a
   ( ) Smartphone/ PDA
   ( ) Cellular phone
   ( ) Touchscreen (e.g. iPhone)

10) What is the brand and model of your mobile device?
    Brand:________________________
    Model:________________________

11) How many text messages do you **send** per week?
    ( ) I send 0-5 text messages per week
    ( ) I send 6-15 text messages per week
    ( ) I send more than 15 text messages per week

12) Do you consider yourself a NOVICE, AVERAGE, OR PROFICIENT communication device user?
    ( ) Novice
    ( ) Average
    ( ) Proficient

13) Which of the following statements best describes your **computer** typing skills?
    ( ) I type using one or two fingers and need to look at the keyboard
    ( ) I type using all fingers and sometimes need to look at the keyboard
    ( ) I type using all fingers without looking at the keyboard

14) What is the one thing you **LIKE** most about your current communication device?

____

15) What is the one thing you **DISLIKE** most about your current communication device?

____

16) Have you ever participated in a consumer test or a focus group about a telecommunications product or service?
    ( ) Yes
    ( ) No

    If YES, to question 16, when was the last time you participated?
    ( ) less than 1 year
    ( ) 1 year or more

Appendix B. Orientation Script

Hi, my name is Barbara. I will be working with you in today's session. Let me explain why we've asked you to come in today.

We are here to test how easy it is to use different keyboard designs. The goal is to get a better understanding of the advantages and disadvantages of the different designs.

You will be performing some typing tasks with each of these devices, and I would like for you to perform as you normally would. For example, try to work at the same speed and with the same attention to detail that you normally do. Overall, do your best and don't be all that concerned with results.

During today's session, I'll be asking you to complete some forms and answer some questions. It is important that you answer truthfully. My only role here today is to discover both the flaws and advantages of these keyboards from your perspective. So do not answer questions based on what you think I want to hear, but rather answer the questions based on what you really think and feel.

You may ask questions at any time, but I may not answer them, since this is a study of the keyboards and we need to see how they work with a person such as yourself working independently.

While you are working, I will be sitting nearby taking some notes and timings. In addition, the session will be videotaped for subsequent analysis. Do you have any questions?

If not, then let's begin by having you sign the informed consent agreement and permission to video record form.

Appendix C. Text Phrases and Analyses of Letter Frequencies

C.1 Experiment #2: Novice Performance

C.1.1 Phrases

1. Mercifully, it was still open.
2. www.flickr.com/explore/
3. www.digg.com/about/
4. 6129 Lees Pike, 317, Falls Church, VA 22041, jadams@aol.com
5. www.yelp.com/miami
6. 36 Amber Dr, Pittsford, NY 14534, ravi.adapathya@kodak.com
7. It ran until past one o'clock.
8. the laser printer is jammed
9. do not feel too bad about it
10. I didn't understand why, Clay.
11. where can my little dog be
12. 311 Wembley Rd, Reisterstown, MD 21136, yxaio@umaryland.edu
13. seasoned golfers love the game
14. 1207 Palo Verde Rd, Irvine, CA 91617, mail@kowym.com
15. 17 Aviation Dr, Winter Haven, FL 33881, jdkochan@aol.com
16. 2 Talbot Pl, Huntington Station, NY 11746, rgulota@tufts.edu
17. How'd you hear about this one?
18. www.yahoo.com/finance
19. miami.craigslist.org/mdc/
20. a big scratch on the tabletop
21. Oh, that's all right, he said.
22. I'll be waiting for you there.
23. www.travelocity.com/vaca23
24. never mix religion and politics
25. www.espn.com/nfl
26. nothing finer than discovering a treasure
27. You are all right, my brother?
28. the kids are very excited
29. I'll just leave the door open.
30. 3320 E 68th Ct, Indianapolis, IN 46220, bill@wrbaynes.com
31. www.giraffe837.com
32. 5303 Foxridge Dr, 301, Mission, KS 66202, daniel@gmail.com
33. yes you are very smart
34. 5825 Tree Line Dr, Madison, WI 53711, gv@trace.wisc.edu
35. No, Cady, he made second team.
36. www.wikipedia.org/wiki/Asia

C.1.2 Analysis of the Letter Frequencies

Analysis of the letter frequencies of the phrases was conducted using a tool developed by Dr. I. Scott MacKenzie – available for download from http://www.yorku.ca/mack/phrasesets.zip).

C.1.2.1 All Phrases

----
phrases: 36
minimum length: 16
maximum length: 60
average phrase length: 34.36
----
words: 182
unique words: 158
minimum length: 1
maximum length: 27
average word length: 5.99
words containing non-letters: 78
----
letters: 1237
correlation with English: 0.8852

C.1.2.2 Words/Spaces and Sentences only

----
phrases: 19
minimum length: 22
maximum length: 43
average phrase length: 30.11
----
words: 111
unique words: 89
minimum length: 1
maximum length: 11
average word length: 4.32
words containing non-letters: 21
----
letters: 572
correlation with English: 0.9552

C.2 Experiment #3: Expert Performance

C.2.1 Phrases

1. It was late, we were playing kissing games, and Jessica and I were called on to kiss in front of the others.
2. He declared the government is thinking of asking for foreign troops if the situation worsens.
3. However, there are still several types of calls that necessitate the use of telephone operators.
4. Investors studying the toll-road bonds for opportunities find that not all roads are nearing their goals.
5. For the year to date, sales of the company's farm equipment dealers still lag about 5% behind 1960.
6. There is so far no evidence to indicate conclusively that this coupling is under enzymatic control.
7. When Huff attempted to cash another $100 check there Monday, hotel officials called police.
8. As soon as you find out if they are Geely and Harris, come on around to the lounge where I'll be waiting.
9. Lucy drew out the chair and sat down; she relaxed a little, and some of the tension went out of her.
10. Both men knew it was in the Norberg family holdings, but to which of the cousins did it belong, Anta or Freya?
11. Matsuo had faked, both the burned and the unburned, the latter decomposing rapidly under the tropical sun.
12. While I respect his sincere concern for peace, he made four points that I would like to question.
13. A body of water is usually the center of interest at parks which attract the greatest picnic and camping use.
14. There was evident delight on the part of the subject in response to her experience of the freedom of movement.
15. By political, economic, geographic and natural standards, they were justified in doing so.
16. At one astronomical unit from the sun (the Earth's distance) the dust orbits are probably nearly circular.
17. Russia's young gymnasts have studied dance before having the rigorous training on apparatus.
18. But this was not unusual, because youth in these quarters was always pushed at a distance from its elders.
19. Johnny vigorously pounded two bleached steer bones against the gourd which served as his drum.
20. It is proposed that in the future complete sampling censuses be carried out at five-year intervals.
21. In most Western cultures today these twins have been sent away to the libraries and museums.

22. Apart from journalese and vaudeville gags, the anatomical is also found in jocular literature.
23. Among the spectators was the noted exotic dancer, Patti Waggin who is Mrs. Don Rudolph when off the stage.
24. Their appearance, next spring, coincides in an almost uncanny way with the flowering of their host plants.
25. The stress on have, which here represents have finished reading the paper, is quite strong.
26. Chandler left Carroll at the bottom of the hill to direct any reinforcements he could find to the fight.
27. There was a very old man and a young woman and a brood of children ranging from toddlers to teen-agers.
28. But don't tell that to a veteran of the Fighting Seventh, especially in a saloon on Saturday night.
29. It squatted low and square upon the sidewalk with a heavy iron grating supporting a glass facade.
30. The ordinance would increase fees from $1 for males and $2 for females to a flat $5 a dog.
31. But remember this - it isn't the aircraft which is vulnerable to nuclear rockets, it is the airfield.
32. Similar payroll tax boosts would be imposed on those under the railroad retirement system.
33. If we return to them today, we have no difficulty spotting their weaknesses but we find them still pleasing.
34. The highroad, one might say at first, belongs to life, while the way to the churchyard belongs to death.
35. In the average situation about one-third of those visited make commitments to Christ and the Church.
36. I have done everything, he wrote, to break up the whole of that unfortunate establishment.
37. The pace could now be accelerated, for the inhabitants of the Aegean stood on firm ground.
38. To my immense relief, she changed the subject in the next sentence: shall we go to the Louvre today?
39. He pulled it over, climbed up, and lifted out the big volume, almost losing his balance from the weight of it.
40. It was the best he could hope for on a watch that had ended with a session in Killpath's office.
41. The Cunard line has under consideration replacing the Queen Mary with a ship smaller than 75,000 tons.
42. Winston took out a pencil, admired the point, and wrote slowly and heavily, clothes stand.
43. Be sure that the landing foot is brought close to the hands and that only one foot lands at a time.
44. Not necessarily to be off all by ourselves, but away from the crowds and common happenstance.

45. A black, snake-like object swayed eerily in front of him, spewing bubbles from its flat cobra head.
46. No amount of religious ceremonies or even joining a church will relieve the gnawing of your inner space.
47. This, together with a derby hat and horn-rim eyeglasses, gave me the appearance of a Russian nihilist.
48. So each reading can be given a weight and each reading a score by adding up these weights.
49. The reality of spirit emerges in this play in spite of the author's convictions to the contrary.
50. John Heffernan, playing Larry Larkin, the cartoonist, carries the show in marvelous fashion.
51. How to feed: for prevention of ketosis, feed 1/4 pound per head daily for 6 weeks commencing at calving time.
52. From 1 July 1958 to 30 June 1960, 24 numbers of the Journal and nine of the Bulletin were published.
53. But he decided he wouldn't mind company in return for free drinks, even though he made good money at his job.
54. In front of you is the Palazzo Madama, once belonging to the Medici and now the Italian Senate.
55. Understanding a work of art involves recognition of the ideas that it reflects or embodies.
56. That is an evening of music-making that would faze many a younger man; Mr. Elman is 70 years old.
57. The third name was (John) Ravencroft, who was admitted to the Inner Temple in November 1631.
58. Sports Writer Ensign Ritchie of the Ogden Standard Examiner went to his compartment to talk with him.
59. It was a real stimulant to a lot of guys I know who have moved past the 2-score-year milestone.
60. Dear sirs: let me begin by clearing up any possible misconception in your minds, wherever you are.
61. He is publicly on record as believing Mr. De Sapio should be replaced for the good of the party.
62. America, America, God shed His grace on thee, and crown thy good with brotherhood from sea to shining sea.
63. Years later, franks-in-buns were accepted as the first to go at the New York Polo Grounds.
64. Work that might cost $500 to $750 in the South could cost $750 to $1,200 in New York City or Chicago.
65. Nevertheless, in another way modern historians still labor in the vineyard of the Oxford school.
66. The Secretary of the Interior may issue rules and regulations to effectuate the purposes of this Act.
67. The moonlit night was made for romance, and he had been looking at her soulfully for some time.

68. Now 38, Mr. Simpkins was graduated from the University of Maryland's College of Agriculture in 1947.
69. But he knew; he sniffed the air and licked it on his lip and knew as a vintner knows a vintage.
70. You could think yourself as grown up as Methuselah, yet the maternal voice still kept its comforting magic.
71. The following discussion of this subject has been adapted from the book Causes Of Catastrophe by L. Don Leet.
72. Art Lund, a fine big actor with a great head of blond hair and a good voice, impersonates Enright.
73. The spirit served chiefly to lull the West while Moscow made inroads into the Middle East.
74. An alternate hatchway entrance, shown on page 25, would reduce the cost of materials $50 to $100.
75. Information on pages 8 to 14 may help you in deciding on the kind and scale of your farming venture.
76. A great deal of labor we have as well, for we are too uncertain of where trust may be placed.
77. Yet paradoxically my liberal friends continue to view Jefferson as one of their patron saints.
78. The window looked out on the Place Redoute - it was the only window of the apartment that did.
79. The illustration (fig. 11) shows this shelter with the roof at ground level and mounded over.
80. During the Civil War, Custer, who achieved a brilliant record, was made brigadier general at the age of 23.
81. I am told that a mortar longer slaked might have remained longer in condition for painting.
82. We had looked forward to what we hoped to be our first informal meeting with a number of Moscow's artists.
83. When cutting the pieces, dress the ends smooth, and square with a smooth file or sanding disk.
84. They hope that if history vouchsafes the West another Budapest, we will receive the opportunity gladly.
85. For you, readers, are an all-important part of the spiritual experiment that is Guideposts.
86. So Prokofieff was able to cultivate his musical talents and harvest a rich reward from them.
87. To the west, the dark green hills of Leyte were lost in the clouds about halfway up their slopes.
88. He was able to find meaning in his art as long as it was the answer to air raids and gas ovens.
89. Experts point to the thinning of pitching talent in the American League caused by expansion.
90. It is most probable that Freud and the Oedipus complex never entered his head in the writing of this story.

91. The fox is all ingratiating smiles when he arrives from New Orleans, accompanied by one wharf rat.
92. On Friday he will go to Portland for the swearing in of Dean Bryson as Multnomah County Circuit Judge.
93. He backed Jess into a corner, grabbed a handful of the man's shirtfront, and drew back his right fist.
94. He saw that Dolores intended to wait until the last minute, thinking he would get nervous.
95. After a while he began to feel better about it, especially when no one bothered to ask any questions.
96. Dams, river development schemes, transportation networks, educational systems require years to construct.
97. I think I would have been much disappointed in Japan if I had not seen Kyoto, Nara, and Hiroshima.
98. One day, to everyone's astonishment, someone drops a match in the powder keg and everything blows up.
99. You can get this added heating feature for as little as $200 more than the price of cooling alone.
100. The rabbi said thoughtfully, I would not want my people to get in trouble with the Church.
101. A supplementary grant from the Geological Society of America helped finance its publication.
102. She seemed so unimpressed that he was obliged to roll up his blue jeans so she could see his brace.
103. I am usually filled with an uneasiness that through some unwitting slip all hell may break loose.
104. He went swiftly up the sidewalk toward the parked car with the two Beach detectives in the front seat.
105. They involve only simple mathematics that are taught in grammar school arithmetic classes.
106. It was mostly for the benefit of the mailman, because hardly anybody else ever visited us.
107. As faulty as has been our leadership clearly the United States must be relied upon to lead.
108. I've never done this before, they always said, waiting for the elevator in the hotel corridor.
109. They remained close together, their air trail wiggling like serpents traveling side by side.
110. It could, by avoiding direct intervention, provide a short-of-war strategy to meet short-of-war infiltration.
111. He expected nothing for himself but that which naturally follows those marked for misfortune.
112. He began to wish that he hadn't shouted that other evening when the truck bore down through the crossing.
113. It is an irritable rule that does baseball more harm than good, especially at the minor league level.

114. If this choice is less exciting than New York Democrats may wish, it nevertheless must be made.
115. I was the first to get my squad on the ball, and anybody thinking it was easy is pretty damn dumb.
116. Since arriving here, however, I have formed a far different religious picture of present-day England.
117. The calibration of piezoelectric sensors in terms of the particle parameters is very uncertain.
118. Dealers would do well to visit such a campground often, look at the equipment and talk with the campers.
119. Such an understanding, although it must seek to be sympathetic, is not a matter of intuition.
120. She refolded the letter, replaced it in its envelope, and turned with relief to one from her brother George.
121. And to offend the dead meant to incur their wrath, and thus provoke the unleashing of countrywide disasters.
122. I knew the only way I could beat you was to play possum, but it was a good try, kid, and I appreciate it.
123. When Alec finished reading he was sure that either Forbes or Stacy had killed Diana Beauclerk.
124. The only evidence of occupation came from the chimney, which was belching out thick smoke.
125. If we did not mean to say this, why should we be so relieved on finding that the suffering had not occurred?
126. And with Progressivism the Religion of Humanity was replacing what Gabriel called Christian supernaturalism.
127. In a lacey open weave shoes have a luster finish, braided collar and bow highlight on the squared throat.
128. The President was even more generous with the First Lady than he had been before the tragedy.
129. Two men, together like us, we could do something fine out there, maybe find a place where no one's ever been.
130. He stood looking down at her for a moment, wondering what could have reduced her to this condition.
131. The other problem is the matter of financing the transition period in the several cities and towns.
132. A study at the Pentagon and at the service academies revealed that nothing was being done there.
133. In the darkness he could see the rosy reflection of the neon sign on the wall opposite the window.
134. The resulting, indescribable torment affects every Southern mind and is the basis of the Southern hysteria.
135. Flushing stadium in works the New York franchise is headed by Mrs. Charles Shipman Payson.
136. The man whose reactions and conclusions get the most space is, of course, the Field Marshal himself.

137. The inadequacy of our library system will become critical unless we act vigorously to correct this condition.
138. In the same way I like to think we owe our loyalty as legislators to our community, our district, our State.
139. Having a boat financed through a local bank is done much the same way as an automobile loan is extended.
140. He had retained his hat and his horn, and, whatever fun might still be going, he was ready to join it.
141. These machines produce the higher quality stretch yarns required in weaving stretch and textured fabrics.
142. In the darkness he could see the rosy reflection of the neon sign on the wall opposite the window.
143. He felt cheerful again, refreshed; presentable in his wide-cut brown suit, the well-made riding boots.
144. The congressman's patriotism is always involved when he turns upon the Defense Department.
145. Independent market owners work six days a week; and my husband hasn't had a vacation in 14 years.
146. He quickly called on Ghana, Tunisia, Morocco, Guinea and Mali, which dispatched troops within hours.
147. The name alkali bee indicates that one has to look for them in rather inhospitable places.
148. Our efforts to overcome the lead of the Russians in space are bound to mean accelerated Federal spending.
149. The unstable political situation there represents one reason new plants shy away from the East Side.
150. They were staring at him in the same blank and menacing way that the men outside the gate had stared.
151. Two minutes later it came again- a double explosion, followed by a third, sounding more distant.
152. Here's a present for you, he said, shoving his bullet-riddled hat down over Nate's purpling forehead.
153. The sun, blazing hot as prophesied, was far from kind to Mrs. Kirby's varicolored properties.
154. As critic Walter Kerr points out: adaptations, so long as they are good, still qualify as creative.
155. He is driven back by his yearning to the wintry homeland of his fathers in the forest of Tiveden.
156. The air took on a special strength now that they'd left the fecund warmth of the farmland behind.
157. Each time a dictionary form matches a text form, the information cell of the matching text form is saved.
158. Serve each breast on a thin slice of slow-baked ham and sprinkle with Thompson seedless grapes.
159. Yet even in these marriages, psychologists say, wives are asserting themselves more strongly.

160. In the last few years the telephone company has managed to automate many areas of their service.
161. It should be enough to say that the practice of the state buying automobiles is at least forty years old.
162. A couple decks for me, Mr. Skyros- and ten-twelve to sell, see, I like to have a little ready cash.
163. He could not recognize it; he was absolutely unfamiliar with it because he had no visual memory at all.
164. Goodwin was telegrapher for the American Telegraph Company and the Troy and Canada Junction Telegraph Company.
165. Are you indiscriminately offering unnecessary medical services- flu shots, sun lamp treatments, etc.?
166. For example: don't wall in your kitchen before you hang the wall cabinets and set the appliances.
167. The building will contain 430,000 square feet, approximately the same as our present plant.
168. That is, they used opaque color throughout, getting solid highlights with active lime white.
169. In 1872 there were known to be twenty-two in Norton County, and one had been in the family for 200 years.
170. It was the hard way to fight a war but Thomas did it without making any disastrous mistakes.
171. In the 1890's the Palace Hotel began serving an oyster dish named after its manager, John C. Kirkpatrick.
172. It was only the other day that I saw something of yours, about something or other, in some magazine.
173. In the casual field straws feature wedge heels of cork or carved wood in a variety of styles.
174. It snowed softly, silently, an undulating interruption of his vision against the night sky.
175. That keeps in the cold, retains moisture and prevents the heaving of alternate freezing and thawing.
176. He keeps riding me because I like to listen to the radio and sing while I'm taking a bath.
177. Within the narrow frame of military tactics, too, the experts agree that the campaign was brilliant.
178. In Nara I stayed at the hotel where the Prince and Princess had stayed on their honeymoon.
179. And Lawrence Chase, son of the Ransom Chases, is listed at his new address in Oxford, England.
180. Of course it was water he really craved; down in the broil of the sun he was becoming dried out.

C.2.2 Analysis of the Letter Frequencies

Analysis of the letter frequencies of the phrases was conducted using a tool developed by Dr. I. Scott MacKenzie – available for download from http://www.yorku.ca/mack/phrasesets.zip).

----
phrases: 180
minimum length: 90
maximum length: 110
average phrase length: 98.61
----
words: 3132
unique words: 1590
minimum length: 1
maximum length: 16
average word length: 4.72
words containing non-letters: 438
----
letters: 17750
correlation with English: 0.9756

Appendix D. XML for Keyboards


```
<?xml version="1.0" encoding="utf-8" ?>
<key>
	<layout1>
		<order order="Space">Space|t|a|w|s|h|f|b|m|c|o|l|p|d|r|n|g|y|e|i|k|u|v|j|q|z|x|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
		<order order="a">Space|n|r|t|s|l|d|i|y|c|v|m|k|b|g|p|u|f|w|x|z|h|j|a|e|o|q|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
		<order order="b">Space|e|u|o|a|l|r|i|j|s|b|y|t|m|v|w|c|d|f|g|h|k|n|p|q|x|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
		<order order="c">Space|h|a|o|e|k|t|l|r|i|u|c|y|s|n|q|b|d|f|g|j|m|p|v|w|x|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
		<order order="d">Space|e|i|o|a|s|r|u|y|d|l|g|v|f|t|n|w|b|j|m|c|h|k|p|q|x|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
		<order order="e">Space|r|n|d|a|s|e|l|t|y|m|c|v|i|w|f|p|x|g|o|k|h|q|b|u|z|j|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
		<order order="f">Space|o|i|r|e|a|t|f|l|u|s|y|m|b|c|d|g|h|j|k|n|p|q|v|w|x|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
		<order order="g">Space|e|h|o|r|i|a|u|l|s|n|g|t|y|b|m|w|c|d|f|j|k|p|q|v|x|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
		<order order="h">Space|e|a|i|o|t|r|u|y|s|w|l|p|h|m|b|c|d|f|g|j|k|n|q|v|x|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
		<order order="i">Space|n|t|s|l|c|r|d|m|g|e|v|o|k|f|p|a|z|x|b|u|q|i|h|j|w|y|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
		<order order="j">Space|u|o|e|i|a|b|c|d|f|g|h|j|k|l|m|n|p|q|r|s|t|v|w|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
		<order order="k">Space|e|i|n|s|l|y|o|u|a|m|p|b|c|d|f|g|h|j|k|q|r|t|v|w|x|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
		<order order="l">Space|e|l|i|o|a|d|y|s|u|t|f|k|p|m|w|v|r|g|c|b|n|h|j|q|x|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
		<order order="m">Space|e|a|o|i|u|p|s|b|m|y|f|l|n|t|c|d|g|h|j|k|q|r|v|w|x|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
		<order order="n">Space|d|g|e|t|o|s|i|y|a|c|l|k|n|u|j|f|h|q|v|b|r|w|x|m|p|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
		<order order="o">Space|u|r|n|m|t|w|o|l|s|p|v|k|d|i|f|b|a|c|y|e|g|z|j|x|h|q|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
		<order order="p">Space|e|l|o|a|r|i|u|p|s|t|h|y|w|c|f|b|d|g|j|k|m|n|q|v|x|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
		<order order="q">Space|u|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|v|w|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
		<order order="r">Space|e|o|s|i|a|t|y|d|n|u|l|r|m|g|k|c|v|h|f|b|p|w|j|q|x|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
		<order order="s">Space|t|e|h|i|o|a|u|s|p|k|l|c|y|w|m|n|b|q|g|d|f|j|r|v|x|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
		<order order="t">Space|h|e|o|a|i|s|r|t|u|y|w|l|c|b|g|n|d|f|m|z|j|k|p|q|v|x|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
		<order order="u">Space|t|r|s|n|l|g|c|m|e|i|b|p|d|a|f|y|o|z|q|x|h|j|k|u|v|w|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
		<order order="v">Space|e|i|o|a|y|u|b|c|d|f|g|h|j|k|l|m|n|p|q|r|s|t|v|w|x|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
		<order order="w">Space|a|h|i|e|o|n|r|s|l|d|t|b|u|k|c|f|g|j|m|p|q|v|w|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
		<order order="x">Space|t|p|a|i|c|e|u|o|b|d|f|g|h|j|k|l|m|n|q|r|s|v|w|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
		<order order="y">Space|o|e|s|i|p|a|b|l|t|w|g|h|m|u|c|d|f|j|k|n|q|r|v|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
		<order order="z">Space|e|z|l|a|s|y|i|o|b|c|d|f|g|h|j|k|m|n|p|q|r|t|u|v|w|x|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
	</layout1>
	<layout2>
		<order order="Space">Caps|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|%|(|)|Enter|0|1|2|3|4|5|6|7|8|9|Space|t|a|w|s|h|f|b|.|,|?|quote|dquote|;|:|!|-|/|@|$</order>
```

```
<order
order="a">Caps|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|%|(|)|Enter|0|1|2|3|4|5|6|7|8|9|Space|n|r|t|s|l|d|i|.|,|?|quote|dquote|;|:|!|-
|/|@|$</order
<order
order="b">Caps|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|%|(|)|Enter|0|1|2|3|4|5|6|7|8|9|Space|e|u|o|a|l|r|i|.|,|?|quote|dquote|;|:|!|-
|/|@|$</order>
<order
order="c">Caps|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|%|(|)|Enter|0|1|2|3|4|5|6|7|8|9|Space|h|a|o|e|k|t|l|.|,|?|quote|dquote|;|:|!|-
|/|@|$</order>
<order
order="d">Caps|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|%|(|)|Enter|0|1|2|3|4|5|6|7|8|9|Space|e|i|o|a|s|r|u|.|,|?|quote|dquote|;|:|!|-
|/|@|$</order>
<order
order="e">Caps|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|%|(|)|Enter|0|1|2|3|4|5|6|7|8|9|Space|r|n|d|a|s|e|l|.|,|?|quote|dquote|;|:|!|-
|/|@|$</order>
<order
order="f">Caps|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|%|(|)|Enter|0|1|2|3|4|5|6|7|8|9|Space|o|i|r|e|a|t|f|.|,|?|quote|dquote|;|:|!|-
|/|@|$</order>
<order
order="g">Caps|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|%|(|)|Enter|0|1|2|3|4|5|6|7|8|9|Space|e|h|o|r|i|a|u|.|,|?|quote|dquote|;|:|!|-
|/|@|$</order>
<order
order="h">Caps|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|%|(|)|Enter|0|1|2|3|4|5|6|7|8|9|Space|e|a|i|o|t|r|u|.|,|?|quote|dquote|;|:|!|-
|/|@|$</order>
<order
order="i">Caps|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|%|(|)|Enter|0|1|2|3|4|5|6|7|8|9|Space|n|t|s|l|c|r|d|.|,|?|quote|dquote|;|:|!|-
|/|@|$</order>
<order
order="j">Caps|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|%|(|)|Enter|0|1|2|3|4|5|6|7|8|9|Space|u|o|e|i|a|b|c|.|,|?|quote|dquote|;|:|!|-
|/|@|$</order>
<order
order="k">Caps|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|%|(|)|Enter|0|1|2|3|4|5|6|7|8|9|Space|e|i|n|s|l|y|o|.|,|?|quote|dquote|;|:|!|-
|/|@|$</order>
<order
order="l">Caps|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|%|(|)|Enter|0|1|2|3|4|5|6|7|8|9|Space|e|l|i|o|a|d|y|.|,|?|quote|dquote|;|:|!|-
|/|@|$</order>
<order
order="m">Caps|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|%|(|)|Enter|0|1|2|3|4|5|6|7|8|9|Space|e|a|o|i|u|p|s|.|,|?|quote|dquote|;|:|!|-
|/|@|$</order>
<order
order="n">Caps|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|%|(|)|Enter|0|1|2|3|4|5|6|7|8|9|Space|d|g|e|t|o|s|i|.|,|?|quote|dquote|;|:|!|-
|/|@|$</order>
<order
order="o">Caps|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|%|(|)|Enter|0|1|2|3|4|5|6|7|8|9|Space|u|r|n|m|t|w|o|.|,|?|quote|dquote|;|:|!|-
|/|@|$</order>
<order
order="p">Caps|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|%|(|)|Enter|0|1|2|3|4|5|6|7|8|9|Space|e|l|o|a|r|i|u|.|,|?|quote|dquote|;|:|!|-
|/|@|$</order>
<order
order="q">Caps|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|%|(|)|Enter|0|1|2|3|4|5|6|7|8|9|Space|u|a|b|c|d|e|f|.|,|?|quote|dquote|;|:|!|-
|/|@|$</order>
<order
order="r">Caps|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|%|(|)|Enter|0|1|2|3|4|5|6|7|8|9|Space|e|o|s|i|a|t|y|.|,|?|quote|dquote|;|:|!|-
|/|@|$</order>
<order
order="s">Caps|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|%|(|)|Enter|0|1|2|3|4|5|6|7|8|9|Space|t|e|h|i|o|a|u|.|,|?|quote|dquote|;|:|!|-
|/|@|$</order>
<order
order="t">Caps|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|%|(|)|Enter|0|1|2|3|4|5|6|7|8|9|Space|h|e|o|a|i|s|r|.|,|?|quote|dquote|;|:|!|-
|/|@|$</order>
<order
order="u">Caps|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|%|(|)|Enter|0|1|2|3|4|5|6|7|8|9|Space|t|r|s|n|l|g|c|.|,|?|quote|dquote|;|:|!|-
|/|@|$</order>
<order
order="v">Caps|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|%|(|)|Enter|0|1|2|3|4|5|6|7|8|9|Space|e|i|o|a|y|u|b|.|,|?|quote|dquote|;|:|!|-
|/|@|$</order>
<order
order="w">Caps|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|%|(|)|Enter|0|1|2|3|4|5|6|7|8|9|Space|a|h|i|e|o|n|r|.|,|?|quote|dquote|;|:|!|-
|/|@|$</order>
```

<order order="x">Caps|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|%|(|)|Enter|0|1|2|3|4|5|6|7|8|9|Space|t|p|a|i|c|e|u|.|,|?|quote|dquote|;|:|!|-|/|@|$</order>
<order order="y">Caps|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|%|(|)|Enter|0|1|2|3|4|5|6|7|8|9|Space|o|e|s|i|p|a|b|.|,|?|quote|dquote|;|:|!|-|/|@|$</order>
<order order="z">Caps|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|%|(|)|Enter|0|1|2|3|4|5|6|7|8|9|Space|e|z|l|a|s|y|i|.|,|?|quote|dquote|;|:|!|-|/|@|$</order>
</layout2>
<layout3>
<order order="Space">Space|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
<order order="a">Space|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
<order order="b">Space|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
<order order="c">Space|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
<order order="d">Space|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
<order order="e">Space|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
<order order="f">Space|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
<order order="g">Space|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
<order order="h">Space|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
<order order="i">Space|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
<order order="j">Space|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
<order order="k">Space|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
<order order="l">Space|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
<order order="m">Space|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
<order order="n">Space|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
<order order="o">Space|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
<order order="p">Space|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
<order order="q">Space|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
<order order="r">Space|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
<order order="s">Space|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
<order order="t">Space|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
<order order="u">Space|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
<order order="v">Space|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
<order order="w">Space|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
<order order="x">Space|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
<order order="y">Space|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
<order order="z">Space|a|b|c|d|e|f|g|h|i|j|k|l|m|n|o|p|q|r|s|t|u|v|w|x|y|z|Caps|0|1|2|3|4|5|6|7|8|9|.|,|?|quote|dquote|;|:|!|-|/|@|$|%|(|)|Enter</order>
</layout3>
<layout4>
<order order="0">0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0|0</order>
</layout4> <key>

Appendix E. XML for Sample Test File


```xml
<?xml version="1.0" encoding="utf-8" ?>
<test>
        <set1>
                <keyboard_row>2</keyboard_row>
                <keyboard_type>1</keyboard_type>
                <phrase>The resulting, indescribable torment affects every Southern mind and is the basis of the Southern
                hysteria.</phrase>
                <hiddendata>FOCL</hiddendata>
                <keyboard_home_location>1</keyboard_home_location>
                <keys_per_row>27</keys_per_row>
        </set1>
        <set2>
                <keyboard_row>2</keyboard_row>
                <keyboard_type>1</keyboard_type>
                <phrase>Flushing stadium in works the New York franchise is headed by Mrs. Charles Shipman
                Payson.</phrase>
                <hiddendata>FOCL</hiddendata>
                <keyboard_home_location>1</keyboard_home_location>
                <keys_per_row>27</keys_per_row>
        </set2>
        <set3>
                <keyboard_row>2</keyboard_row>
                <keyboard_type>1</keyboard_type>
                <phrase>In the darkness he could see the rosy reflection of the neon sign on the wall opposite the
                window.</phrase>
                <hiddendata>FOCL</hiddendata>
                <keyboard_home_location>1</keyboard_home_location>
                <keys_per_row>27</keys_per_row>
        </set3>
        <set4>
                <keyboard_row>1</keyboard_row>
                <keyboard_type>4</keyboard_type>
                <phrase>STOP! Please complete our questionnaire at this time.  Thank You!</phrase>
                <hiddendata>Questionnaire Prompt</hiddendata>
                <keyboard_home_location>1</keyboard_home_location>
        </set4>
        <set5>
                <keyboard_row>2</keyboard_row>
                <keyboard_type>3</keyboard_type>
                <phrase>The other problem is the matter of financing the transition period in the several cities and
                towns.</phrase>
                <hiddendata>ABC</hiddendata>
                <keyboard_home_location>1</keyboard_home_location>
                <keys_per_row>27</keys_per_row>
        </set5>
        <set6>
                <keyboard_row>2</keyboard_row>
                <keyboard_type>3</keyboard_type>
                <phrase>He stood looking down at her for a moment, wondering what could have reduced her to this
                condition.</phrase>
                <hiddendata>ABC</hiddendata>
                <keyboard_home_location>1</keyboard_home_location>
                <keys_per_row>27</keys_per_row>
        </set6>
        <set7>
                <keyboard_row>2</keyboard_row>
                <keyboard_type>3</keyboard_type>
                <phrase>The President was even more generous with the First Lady than he had been before the
                tragedy.</phrase>
                <hiddendata>ABC</hiddendata>
                <keyboard_home_location>1</keyboard_home_location>
                <keys_per_row>27</keys_per_row>
        </set7>
        <set8>
                <keyboard_row>1</keyboard_row>
                <keyboard_type>4</keyboard_type>
```

<phrase>STOP! Please complete our questionnaire at this time. Thank You!</phrase>
<hiddendata>Questionnaire Prompt</hiddendata>
<keyboard_home_location>1</keyboard_home_location>
</set8>
<set9>
<keyboard_row>2</keyboard_row>
<keyboard_type>2</keyboard_type>
<phrase>A study at the Pentagon and at the service academies revealed that nothing was being done there.</phrase>
<hiddendata>HYBRID</hiddendata>
<keyboard_home_location>42</keyboard_home_location>
<keys_per_row>31</keys_per_row>
</set9>
<set10>
<keyboard_row>2</keyboard_row>
<keyboard_type>2</keyboard_type>
<phrase>Two men, together like us, we could do something fine out there, maybe find a place where no one's ever been.</phrase>
<hiddendata>HYBRID</hiddendata>
<keyboard_home_location>42</keyboard_home_location>
<keys_per_row>31</keys_per_row>
</set10>
<set11>
<keyboard_row>2</keyboard_row>
<keyboard_type>2</keyboard_type>
<phrase>In a lacey open weave shoes have a luster finish, braided collar and bow highlight on the squared throat.</phrase>
<hiddendata>HYBRID</hiddendata>
<keyboard_home_location>42</keyboard_home_location>
<keys_per_row>31</keys_per_row>
</set11>
<set12>
<keyboard_row>1</keyboard_row>
<keyboard_type>4</keyboard_type>
<phrase>STOP! You are almost done. All you have to do now is complete a few more questionnaires. Thank You!</phrase>
<hiddendata>STOP</hiddendata>
<keyboard_home_location>1</keyboard_home_location>
</set12>
</test>

Appendix F. Immediate Usability Results[6]

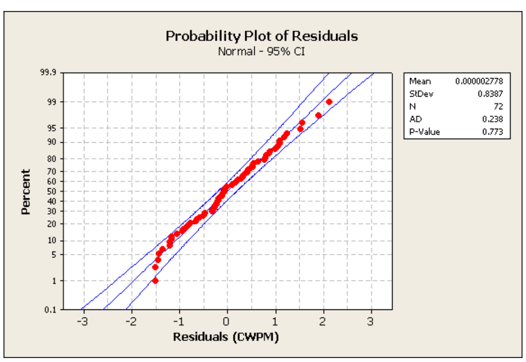
F.1 CWPM

F.1.1 CWPM ANOVA Results (Full Factorial Model)

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	16	697.727	.000
Age	1	16	14.366	.002
Gender	1	16	.017	.896
Experience (EXP)	1	16	.011	.918
Keyboard	2	32.000	.953	.396
Age * Gender	1	16	.496	.492
Age * EXP	1	16	.030	.865
Age * Keyboard	2	32.000	.058	.944
Gender * EXP	1	16	.514	.484
Gender * Keyboard	2	32.000	2.800	.076
EXP * Keyboard	2	32.000	.241	.787
Age * Gender * EXP	1	16	.011	.918
Age * Gender * Keyboard	2	32.000	1.705	.198
Age * EXP * Keyboard	2	32.000	.164	.850
Gender * EXP * Keyboard	2	32.000	1.010	.376
Age * Gender * EXP * Keyboard	2	32.000	.809	.454

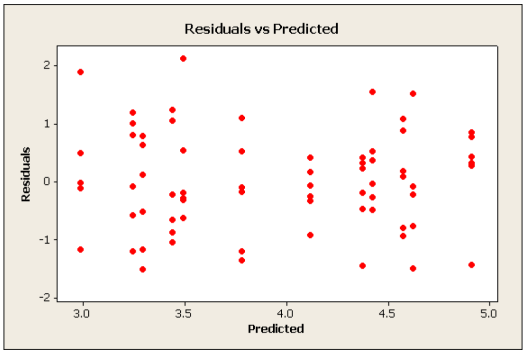
F.1.2 CWPM Test of Model Assumptions for Final Model

F.1.2.1 Test of Normality

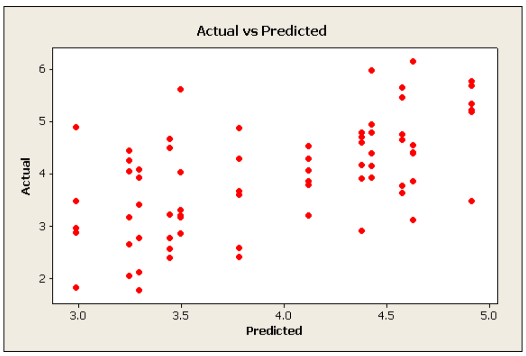


6 ABC= Alphabetical, FOCL= Predictive, HYB= Hybrid

F.1.2.2 Test of Homogeneity of Variance



F.1.2.3 Goodness-of-fit



F.1.3 CWPM Correlations

	ABC CWPM	FOCL CWPM
FOCL CWPM	0.352 *0.092*	
HYB CWPM	0.493 *0.014*	0.629 *0.001*

Cell Contents: Pearson correlation
P-Value

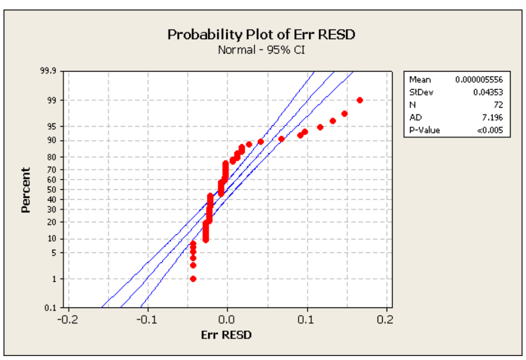
F.2 Uncorrected Error Rate

F.2.1 UER ANOVA Results (Full Factorial Model)

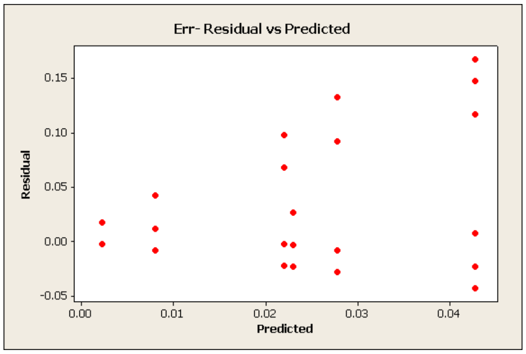
Source	Numerator df	Denominator df	F	Sig.
Intercept	1	16	14.541	.002
Age	1	16	5.756	.029
Gender	1	16	.695	.417
EXP	1	16	2.373	.143
Keyboard	2	32.000	1.159	.327
Age * Gender	1	16	.108	.747
Age * EXP	1	16	2.072	.169
Age * Keyboard	2	32.000	.845	.439
Gender * EXP	1	16	.052	.823
Gender * Keyboard	2	32.000	.341	.713
EXP * Keyboard	2	32.000	.044	.957
Age * Gender * EXP	1	16	.695	.417
Age * Gender * Keyboard	2	32.000	.114	.893
Age * EXP * Keyboard	2	32.000	.341	.713
Gender * EXP * Keyboard	2	32.000	.063	.939
Age * Gender * EXP * Keyboard	2	32.000	.195	.824

F.2.2 UER Test of Model Assumptions for Final Model

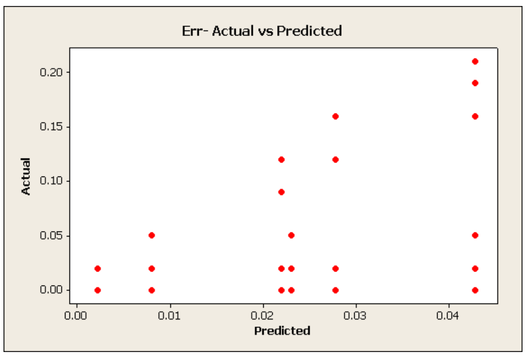
F.2.2.1 Test of Normality



F.2.2.2 Test of Homogeneity of Variance



F.2.2.3 Goodness-of-fit



F.2.3 UER Friedman's Test

Friedman Test: UER vs. Keyboard Blocked by Participant

S = 1.40 DF = 2 P = 0.498
S = 2.68 DF = 2 P = 0.262 (adjusted for ties)

Keyboard	N	Est Median	Sum of Ranks
ABC	24	0.00000	44.5
FOCL	24	0.00000	47.0
HYBRID	24	0.00000	52.5

Grand median = 0.00000

F.2.4 UER Logistic Regression

Tests of Model Effects

Source	Type III		
	Wald Chi-Square	df	Sig.
(Intercept)	7.588	1	.006
Gender	3.727	1	.054
Keyboard	1.339	2	.512

Dependent Variable: UER
Model: (Intercept), Gender, Keyboard

Parameter Estimates

Parameter	B	Std. Error	95% Wald Confidence Interval		Hypothesis Test		
			Lower	Upper	Wald Chi-Square	df	Sig.
(Intercept)	-.887	.5034	-1.873	.100	3.104	1	.078
[Gender=Female]	1.054	.5457	-.016	2.123	3.727	1	.054
[Gender=Male]	0	.	.	.	.	.	.
[Keyboard=Alphabetical]	-.808	.6989	-2.178	.562	1.337	1	.248
[Keyboard=Predictive]	-.380	.6015	-1.559	.799	.399	1	.528
[Keyboard=Hybrid]	0	.	.	.	.	.	.
(Scale)	1						

Dependent Variable: UER
Model: (Intercept), Gender, Keyboard

F.3 KSPC

F.3.1 KSPC Correlations

	ABC KSPC	FOCL KSPC
FOCL KSPC	0.143 *0.506*	
HYB KSPC	0.081 *0.705*	0.085 *0.692*

Cell Contents: Pearson correlation
P-Value

F.3.2 KSPC ANOVA Results (Full Factorial Model)

Type III Tests of Fixed Effects[a]

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	16.000	1781.503	.000
Age	1	16.000	.194	.665
Gender	1	16.000	.041	.842
EXP	1	16.000	.328	.575
Keyboard	2	32.000	46.453	.000
Age * Gender	1	16.000	.422	.525
Age * EXP	1	16.000	4.952	.041
Age * Keyboard	2	32.000	1.546	.229
Gender * EXP	1	16.000	.027	.871
Gender * Keyboard	2	32.000	3.712	.036
EXP * Keyboard	2	32.000	.659	.524
Age * Gender * EXP	1	16.000	3.641	.074
Age * Gender * Keyboard	2	32.000	.582	.565
Age * EXP * Keyboard	2	32.000	.019	.981
Gender * EXP * Keyboard	2	32.000	1.077	.353
Age * Gender * EXP * Keyboard	2	32.000	1.191	.317

a. Dependent Variable: KSPC.

F.3.3 KSPC ANOVA Results on Keyboard, Age, Gender, Experience, Age by Experience, Keyboard by Gender

Type III Tests of Fixed Effects[a]

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	19	1683.281	.000
Age	1	19	.184	.673
Gender	1	19	.038	.847
Keyboard	2	44	48.481	.000
Gender * Keyboard	2	44	3.876	.028
Experience	1	19	.307	.586
Age * Experience	1	19	4.676	.044

a. Dependent Variable: KSPC.

Pairwise Comparisons[b]

(I) Keyboard	(J) Keyboard	Mean Difference (I-J)	Std. Error	df	Sig.[a]	95% Confidence Interval for Difference[a] Lower Bound	Upper Bound
ABC_CWPM	FOCL_CWPM	3.566*	.468	44	.000	2.624	4.509
	HYB_CWPM	4.307*	.468	44	.000	3.364	5.250
FOCL_CWPM	ABC_CWPM	-3.566*	.468	44	.000	-4.509	-2.624
	HYB_CWPM	.741	.468	44	.120	-.202	1.683
HYB_CWPM	ABC_CWPM	-4.307*	.468	44	.000	-5.250	-3.364
	FOCL_CWPM	-.741	.468	44	.120	-1.683	.202

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

b. Dependent Variable: KSPC.

Pairwise Comparisons[b]

Keyboard	(I) Gender	(J) Gender	Mean Difference (I-J)	Std. Error	df	Sig.[a]	95% Confidence Interval for Difference[a] Lower Bound	Upper Bound
ABC_CWPM	Female	Male	1.551*	.706	59.425	.032	.139	2.963
	Male	Female	-1.551*	.706	59.425	.032	-2.963	-.139
FOCL_CWPM	Female	Male	-.947	.706	59.425	.185	-2.358	.465
	Male	Female	.947	.706	59.425	.185	-.465	2.358
HYB_CWPM	Female	Male	-.338	.706	59.425	.634	-1.750	1.074
	Male	Female	.338	.706	59.425	.634	-1.074	1.750

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

b. Dependent Variable: KSPC.

Pairwise Comparisons[b]

Gender	(I) Keyboard	(J) Keyboard	Mean Difference (I-J)	Std. Error	df	Sig.[a]	95% Confidence Interval for Difference[a] Lower Bound	Upper Bound
Female	ABC_CWPM	FOCL_CWPM	4.815*	.661	44	.000	3.482	6.148
		HYB_CWPM	5.251*	.661	44	.000	3.918	6.585
	FOCL_CWPM	ABC_CWPM	-4.815*	.661	44	.000	-6.148	-3.482
		HYB_CWPM	.436	.661	44	.513	-.897	1.770
	HYB_CWPM	ABC_CWPM	-5.251*	.661	44	.000	-6.585	-3.918
		FOCL_CWPM	-.436	.661	44	.513	-1.770	.897
Male	ABC_CWPM	FOCL_CWPM	2.318*	.661	44	.001	.985	3.651
		HYB_CWPM	3.363*	.661	44	.000	2.030	4.696
	FOCL_CWPM	ABC_CWPM	-2.318*	.661	44	.001	-3.651	-.985
		HYB_CWPM	1.045	.661	44	.121	-.288	2.378
	HYB_CWPM	ABC_CWPM	-3.363*	.661	44	.000	-4.696	-2.030
		FOCL_CWPM	-1.045	.661	44	.121	-2.378	.288

Based on estimated marginal means

*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

Pairwise Comparisons[b]

Gender	(I) Keyboard	(J) Keyboard	Mean Difference (I-J)	Std. Error	df	Sig.[a]	95% Confidence Interval for Difference[a] Lower Bound	Upper Bound
Female	ABC_CWPM	FOCL_CWPM	4.815*	.661	44	.000	3.482	6.148
		HYB_CWPM	5.251*	.661	44	.000	3.918	6.585
	FOCL_CWPM	ABC_CWPM	-4.815*	.661	44	.000	-6.148	-3.482
		HYB_CWPM	.436	.661	44	.513	-.897	1.770
	HYB_CWPM	ABC_CWPM	-5.251*	.661	44	.000	-6.585	-3.918
		FOCL_CWPM	-.436	.661	44	.513	-1.770	.897
Male	ABC_CWPM	FOCL_CWPM	2.318*	.661	44	.001	.985	3.651
		HYB_CWPM	3.363*	.661	44	.000	2.030	4.696
	FOCL_CWPM	ABC_CWPM	-2.318*	.661	44	.001	-3.651	-.985
		HYB_CWPM	1.045	.661	44	.121	-.288	2.378
	HYB_CWPM	ABC_CWPM	-3.363*	.661	44	.000	-4.696	-2.030
		FOCL_CWPM	-1.045	.661	44	.121	-2.378	.288

Based on estimated marginal means

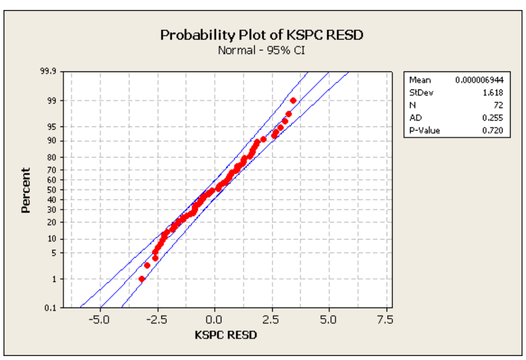
*. The mean difference is significant at the .05 level.

a. Adjustment for multiple comparisons: Least Significant Difference (equivalent to no adjustments).

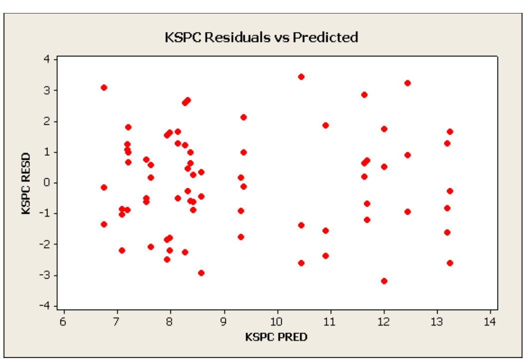
b. Dependent Variable: KSPC.

F.3.4 KSPC Test of Model Assumptions for Final Model

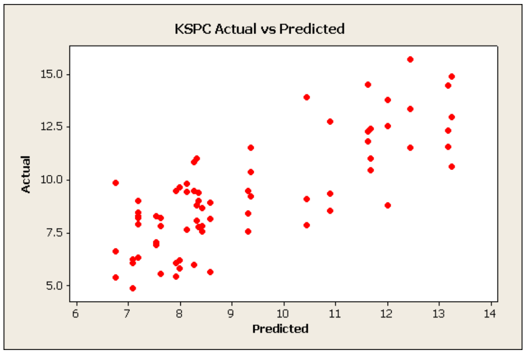
F.3.4.1 Test of Normality



F.3.4.2 Test of Homogeneity of Variance



F.3.4.3.Goodness-of-fit



F.4 Typematic Events

F.4.1 TE ANOVA Results (Full Factorial Model)

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	16	145.912	.000
Age	1	16	.189	.670
Gender	1	16	.431	.521
EXP	1	16	.006	.940
Keyboard	2	32.000	17.304	.000
Age * Gender	1	16	.021	.887
Age * EXP	1	16	.012	.913
Age * Keyboard	2	32.000	2.258	.121
Gender * EXP	1	16	.576	.459
Gender * Keyboard	2	32.000	.541	.587
EXP * Keyboard	2	32.000	.572	.570
Age * Gender * EXP	1	16	.576	.459
Age * Gender * Keyboard	2	32.000	.021	.979
Age * EXP * Keyboard	2	32.000	1.854	.173
Gender * EXP * Keyboard	2	32.000	1.439	.252
Age * Gender * EXP * Keyboard	2	32.000	1.100	.345

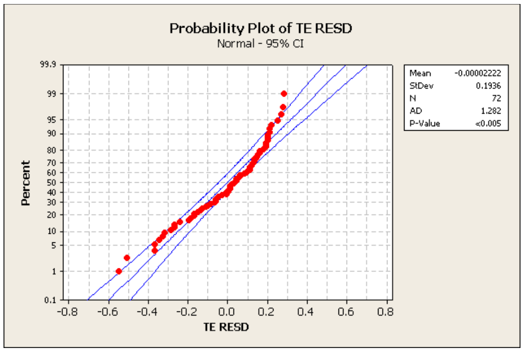
a. Dependent Variable: TE.

F.4.2 TE ANOVA Results on Keyboard for Final Model

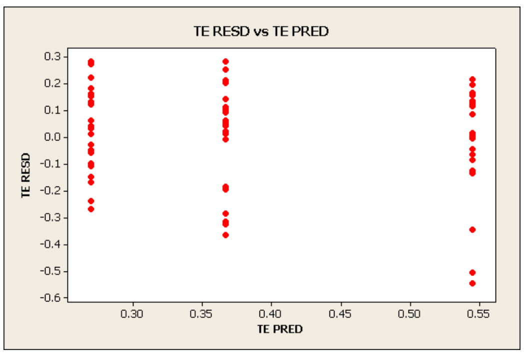
Source	Numerator df	Denominator df	F	Sig.
Intercept	1	23	188.548	.000
Keyboard	2	46	16.726	.000

F.4.3 TE Test of Model Assumptions for Final Model

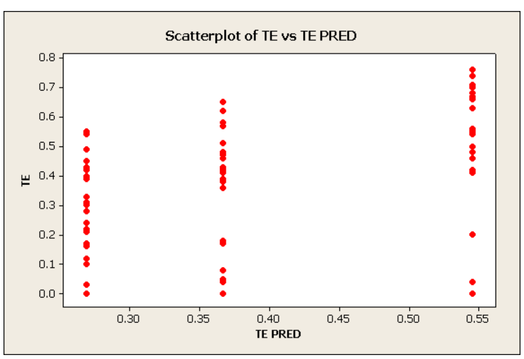
F.4.3.1 Test of Normality



F.4.3.2 Test of Homogeneity of Variance



F.4.3.3 Goodness-of-fit



F.4.4 TE Friedman's Test

Friedman Test: TE versus Keyboard blocked by Subject

S = 20.15 DF = 2 P = 0.000
S = 21.25 DF = 2 P = 0.000 (adjusted for ties)

Keyboard	N	Est Median	Ranks
Alphabetical	24	0.6100	65.0
Predictive	24	0.4133	44.5
Hybrid	24	0.3067	34.5

Grand median = 0.4433

Rank averages and totals

Level	Rank average	Rank total
1: Alphabetical	2.71	65
2: Predictive	1.85	45
3: Hybrid	1.44	35

Friedman's S = 483.5

CHI(2)= 20.15, p=0.0000 (No correction for ties.)
CHI(2)= 21.25, p=0.0000 (With correction for ties.)

Table of rank frequencies

	Conditions		
Rank	1	2	3
1.0	2	6	14
1.5	0	1	1
2.0	3	14	8
2.5	0	0	0
3.0	19	3	1
Rank Tot	65	45	35
Rank Ave	2.7	1.9	1.4

Multiple Comparisons Based on Friedman Rank-Sums

Condition A	Condition B	Absolute Difference	Probability Level
1 : 65	2 : 44.5	20.5	<.01
1 : 65	3 : 34.5	30.5	<.0001
2 : 44.5	3 : 34.5	10	>.20

F.4.5 TE Correlations

	Alphabetical	Predictive
Predictive	0.152	
	0.479	
Hybrid	0.156	0.542
	0.468	*0.006*

Cell Contents: Pearson correlation
P-Value

F.5 Movement Inefficiency

F.5.1 MI Correlations: Keyboards

	FOCL	HYB
HYB	-0.015	
	0.943	
ABC	0.124	0.199
	0.565	*0.352*

Cell Contents: Pearson correlation
P-Value

F.5.2 MI ANOVA Results (Full Factorial Model)

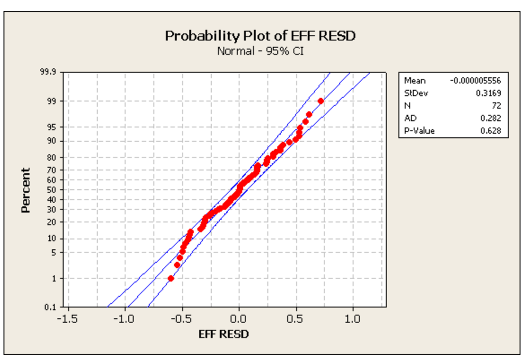
Type III Tests of Fixed Effects[a]

Source	Numerator df	Denominator df	F	Sig.
Intercept	1	16	317.570	.000
Age	1	16	.644	.434
Gender	1	16	.017	.898
EXP	1	16	.302	.590
Keyboard	2	32.000	.463	.634
Age * Gender	1	16	1.040	.323
Age * EXP	1	16	7.593	.014
Age * Keyboard	2	32.000	2.260	.121
Gender * EXP	1	16	.003	.955
Gender * Keyboard	2	32.000	2.565	.093
EXP * Keyboard	2	32.000	.932	.404
Age * Gender * EXP	1	16	4.552	.049
Age * Gender * Keyboard	2	32.000	.945	.399
Age * EXP * Keyboard	2	32.000	.032	.968
Gender * EXP * Keyboard	2	32.000	1.037	.366
Age * Gender * EXP * Keyboard	2	32.000	2.705	.082

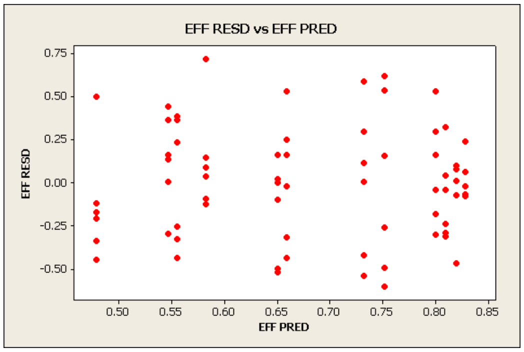
a. Dependent Variable: MI

F.5.3 MI Test of Model Assumptions for Final Model

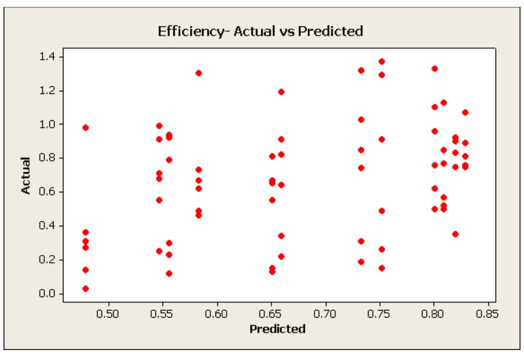
F.5.3.1 Test of Normality



F.5.3.2 Test of Homogeneity of Variance



F.5.3.3 Goodness-of-fit



F.6 Keyboard Ranking

Friedman Test: Rank vs. Keyboard Blocked by Participant

S = 3.08 DF = 2 P = 0.214

Keyboard	N	Est Median	Sum of Ranks
ABC	24	2.0000	44.0
FOCL	24	2.6667	55.0
HYB	24	1.3333	45.0

Grand median = 2.0000

Appendix G. Novice Performance Results[7]

G.1 CWPM

G.1.1 CWPM Correlations

	ABC CWPM	FOCL CWPM
FOCL CWPM	0.480	
	0.000	
HYB CWPM	0.508	0.439
	0.000	*0.000*

Cell Contents: Pearson correlation
P-Value

G.1.2 CWPM T-tests by Keyboard

Variable	N	Mean	StDev	SE Mean	95% Upper Bound	T	P
ABC	288	4.9797	1.2434	0.0733	5.1006	-0.28	0.391
FOCL	288	5.0303	1.5326	0.0903	5.1793	0.34	0.631
HYB	288	5.087	1.716	0.101	5.253	0.86	0.804

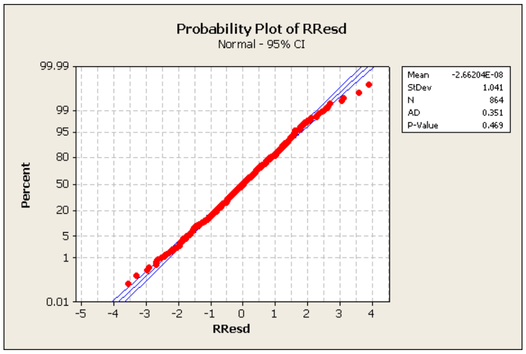
7 ABC= Alphabetical, FOCL= Predictive, HYB= Hybrid

G.1.3 Repeated Measures ANOVA- Full Model

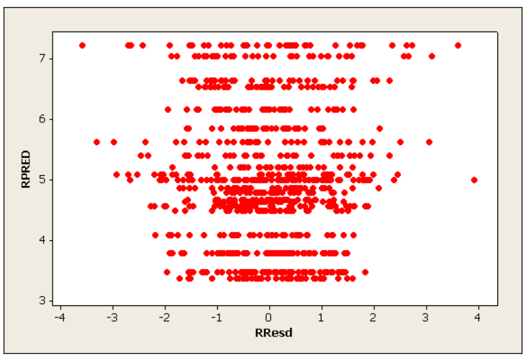
Source	Numerator df	Denominator df	F	Sig.
Intercept	1	16.000	788.108	.000
Age	1	16.000	12.927	.002
Gender	1	16.000	.016	.900
EXP	1	16.000	.000	.998
Type	3	752.000	413.475	.000
Keyboard	2	752.000	1.688	.186
Age * Gender	1	16.000	.356	.559
Age * EXP	1	16.000	.235	.634
Age * Type	3	752.000	5.977	.000
Age * Keyboard	2	752.000	6.030	.003
Gender * EXP	1	16.000	.965	.341
Gender * Type	3	752.000	.026	.994
Gender * Keyboard	2	752.000	.292	.747
EXP * Type	3	752.000	2.137	.094
EXP * Keyboard	2	752.000	3.936	.020
Type * Keyboard	6	752.000	16.669	.000
Age * Gender * EXP	1	16.000	.483	.497
Age * Gender * Type	3	752.000	.352	.788
Age * Gender * Keyboard	2	752.000	.262	.770
Age * EXP * Type	3	752.000	1.041	.374
Age * EXP * Keyboard	2	752.000	4.941	.007
Age * Type * Keyboard	6	752.000	1.818	.093
Gender * EXP * Type	3	752.000	.657	.579
Gender * EXP * Keyboard	2	752.000	8.484	.000
Gender * Type * Keyboard	6	752.000	.853	.529
EXP * Type * Keyboard	6	752.000	1.294	.257
Age * Gender * EXP * Type	3	752.000	1.666	.173
Age * Gender * EXP * Keyboard	2	752.000	5.503	.004
Age * Gender * Type * Keyboard	6	752.000	.390	.886
Age * EXP * Type * Keyboard	6	752.000	4.609	.000
Gender * EXP * Type * Keyboard	6	752.000	1.163	.324
Age * Gender * EXP * Type * Keyboard	6	752.000	1.619	.139

G.1.4 CWPM Test of Model Assumptions for Final Model

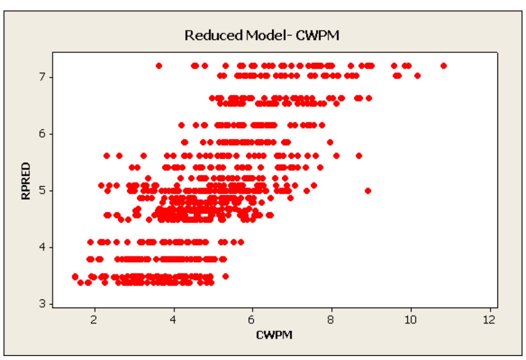
G.1.4.1 Test of Normality



G.1.4.2 Test of Homogeneity of Variance



G.1.4.3 Goodness of Fit



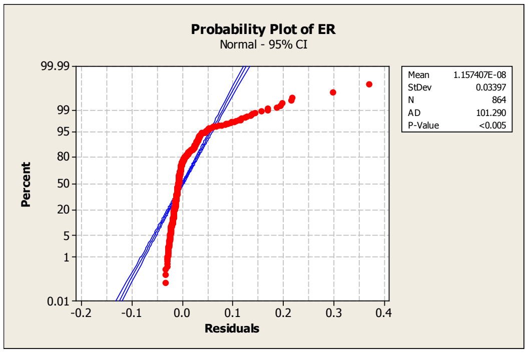
G.2 Uncorrected Error Rate

G.2.1 UER Repeated Measures ANOVA- Full Model

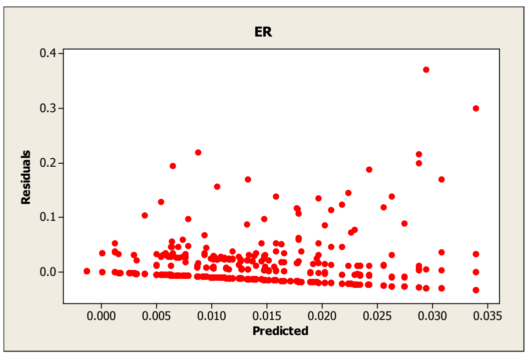
Source	Numerator df	Denominator df	F	Sig.
Intercept	1	16.000	115.770	.000
Age	1	16.000	2.635	.124
Gender	1	16.000	3.676	.073
EXP	1	16.000	2.661	.122
Type	3	752.000	1.866	.134
Keyboard	2	752.000	1.333	.264
Age * Gender	1	16.000	3.538	.078
Age * EXP	1	16.000	.115	.739
Age * Type	3	752.000	3.030	.029
Age * Keyboard	2	752.000	1.131	.323
Gender * EXP	1	16.000	10.128	.006
Gender * Type	3	752.000	.498	.684
Gender * Keyboard	2	752.000	.623	.537
EXP * Type	3	752.000	.233	.873
EXP * Keyboard	2	752.000	.593	.553
Type * Keyboard	6	752.000	.714	.638
Age * Gender * EXP	1	16.000	1.013	.329
Age * Gender * Type	3	752.000	1.535	.204
Age * Gender * Keyboard	2	752.000	.321	.726
Age * EXP * Type	3	752.000	2.225	.084
Age * EXP * Keyboard	2	752.000	.912	.402
Age * Type * Keyboard	6	752.000	.978	.439
Gender * EXP * Type	3	752.000	1.274	.282
Gender * EXP * Keyboard	2	752.000	1.155	.316
Gender * Type * Keyboard	6	752.000	.653	.688
EXP * Type * Keyboard	6	752.000	1.158	.327
Age * Gender * EXP * Type	3	752.000	3.207	.023
Age * Gender * EXP * Keyboard	2	752.000	2.946	.053
Age * Gender * Type * Keyboard	6	752.000	.842	.537
Age * EXP * Type * Keyboard	6	752.000	1.112	.353
Gender * EXP * Type * Keyboard	6	752.000	.588	.740
Age * Gender * EXP * Type * Keyboard	6	752.000	.729	.627

G.2.2 UER Test of Model Assumptions for Final Model

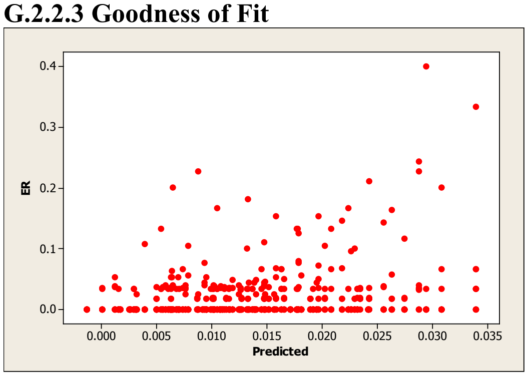
G.2.2.1 Test of Normality



G.2.2.2 Test of Homogeneity of Variance



G.2.2.3 Goodness of Fit



G.2.3 Friedman Tests

Rank averages and totals for Keyboard

Level	Rank average	Rank total
ABC	2.13	51
Predictive	1.98	48
Hybrid	1.90	46

Friedman's S = 15.5

CHI(2)= 0.65, p=0.7240 (No correction for ties.)
CHI(2)= 0.65, p=0.7216 (With correction for ties.)

Rank averages and totals for Text Types

Level	Rank average	Rank total
Address	2.88	69
Sentence	2.79	67
Web	2.13	51
Words	2.21	53

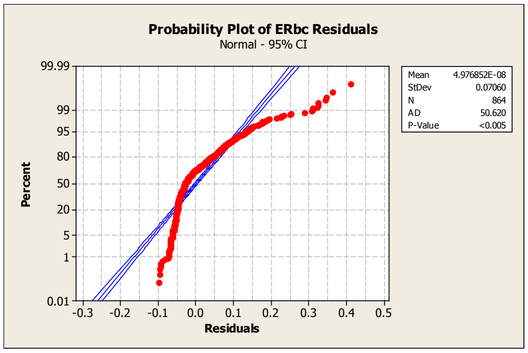
Friedman's S = 260

CHI(3)= 6.50, p=0.0897 (No correction for ties.)
CHI(3)= 6.70, p=0.0823 (With correction for ties.)

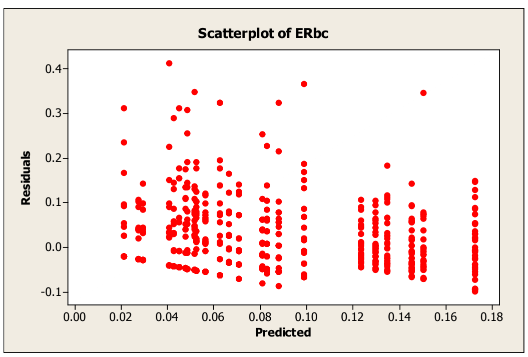
G.3 Total Error Rate

G.3.1 TER Test of Assumptions for Final Model

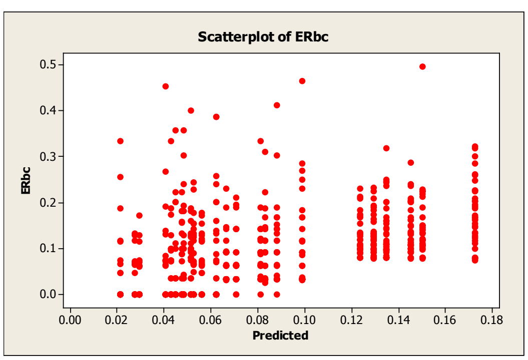
G.3.1.1 Test of Normality



G.3.1.2 Test of Homogeneity of Variance



G.3.1.3 Goodness of Fit



G.3.2 TER Friedman Tests

Rank averages and totals for Text Type

Level	Rank average	Rank total
-----	------------	----------
(1) Addresses	4.00	96
(2) Sentences	2.83	68
(3) Web	1.92	46
(4) Words	1.25	30

Friedman's S = 2456

CHI(3)= 61.40, p=0.0000 (No correction for ties.)
No ties in this data set.
==

Table of rank frequencies for Text Types

	Conditions			

Rank	1	2	3	4
1.0	0	0	6	18
1.5	0	0	0	0
2.0	0	4	14	6
2.5	0	0	0	0
3.0	0	20	4	0
3.5	0	0	0	0
4.0	24	0	0	0
Rank Tot	96	68	46	30
Rank Ave	4.0	2.8	1.9	1.3

==

Multiple Comparisons Based on Friedman Rank-Sums

Condition A	Condition B	Absolute Difference	Probability Level
-----------	-----------	-------------------	-----------------
1 : 96	2 : 68	28	<.01
1 : 96	3 : 46	50	<.0001
1 : 96	4 : 30	66	<.0001
2 : 68	3 : 46	22	<.10
2 : 68	4 : 30	38	<.0005
3 : 46	4 : 30	16	>.20

==

Rank averages and totals for Keyboards

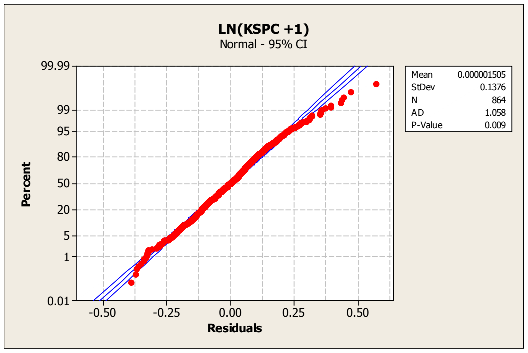
Level	Rank average	Rank total
-----	------------	----------
1: Alphabetical	1.92	46
2: Predictive	1.83	44
3: Hybrid	2.25	54

Friedman's S = 56
CHI(2)= 2.33, p=0.3114 (No correction for ties.)
No ties in this data set.

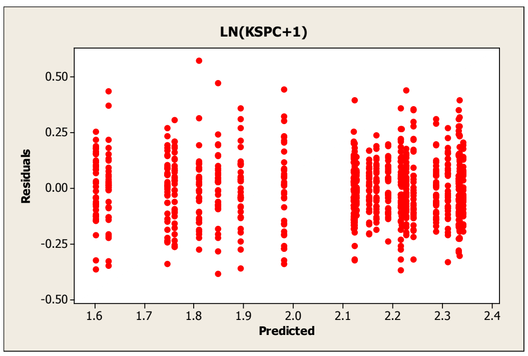
G.4 KSPC

G.4.1 Test of Assumptions for Final Model

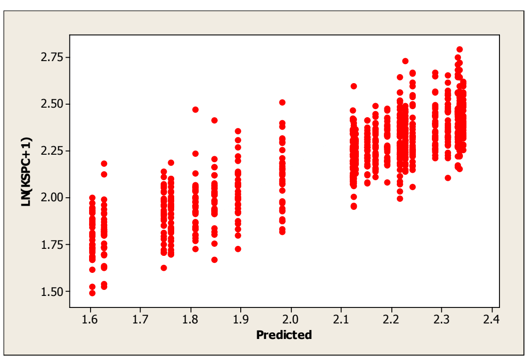
G.4.1.1 Test of Normality



G.4.1.2 Test of Homogeneity of Variance



G.4.1.3 Goodness of Fit



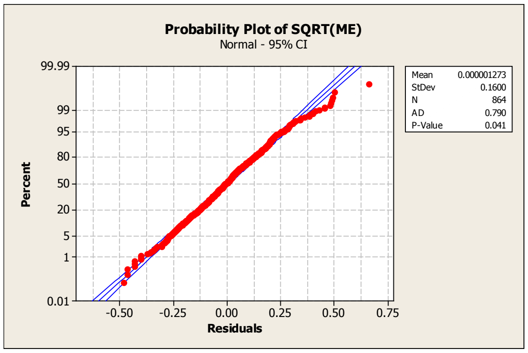
G.5 Movement Inefficiency

G.5.1 MI Computed Minimal Path

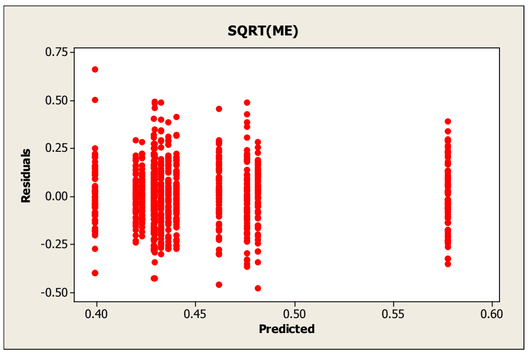
Phrase	*Type*	*Alphabetical*	*Predictive*	*Hybrid*
Mercifully, it was still open.	Sentence	252	180	144
www.flickr.com/explore/	Web	210	152	164
www.digg.com/about/	Web	154	139	140
6129 Lees Pike, 317, Falls Church, VA 22041, jadams@aol.com	Address	442	393	422
www.yelp.com/miami	Web	165	144	143
36 Amber Dr, Pittsford, NY 14534, ravi.adapathya@kodak.com	Address	436	444	432
It ran until past one o'clock.	Sentence	254	191	163
the laser printer is jammed	Words	208	126	90
do not feel too bad about it	Words	186	147	118
I didn't understand why, Clay.	Sentence	221	184	190
where can my little dog be	Words	171	131	109
311 Wembley Rd, Reisterstown, MD 21136, yxaio@umaryland.edu	Address	433	375	410
seasoned golfers love the game	Words	225	158	119
1207 Palo Verde Rd, Irvine, CA 91617, mail@kowym.com	Address	410	357	381
17 Aviation Dr, Winter Haven, FL 33881, jdkochan@aol.com	Address	453	357	372
2 Talbot Pl, Huntington Station, NY 11746, rgulota@tufts.edu	Address	497	399	416
How'd you hear about this one?	Sentence	221	153	146
www.yahoo.com/finance	Web	178	143	162
miami.craigslist.org/mdc/	Web	231	212	160
a big scratch on the tabletop	Words	189	161	132
Oh, that's all right, he said.	Sentence	235	181	153
I'll be waiting for you there.	Sentence	229	127	120
www.travelocity.com/vaca23	Web	189	191	214
never mix religion and politics	Words	258	189	143
www.espn.com/nfl	Web	156	131	125
nothing finer than discovering a treasure	Words	316	170	151
You are all right, my brother?	Sentence	226	159	156
the kids are very excited	Words	149	108	106
I'll just leave the door open.	Sentence	248	159	133
3320 E 68th Ct, Indianapolis, IN 46220, bill@wrbaynes.com	Address	444	372	400
www.giraffe837.com	Web	149	142	122
5303 Foxridge Dr, 301, Mission, KS 66202, daniel@gmail.com	Address	467	397	409
yes you are very smart	Words	131	96	108
5825 Tree Line Dr, Madison, WI 53711, gv@trace.wisc.edu	Address	419	368	394
No, Cady, he made second team.	Sentence	214	179	171
www.wikipedia.org/wiki/Asia	Web	224	207	213

G.5.2 MI Test of Assumptions for Final Model

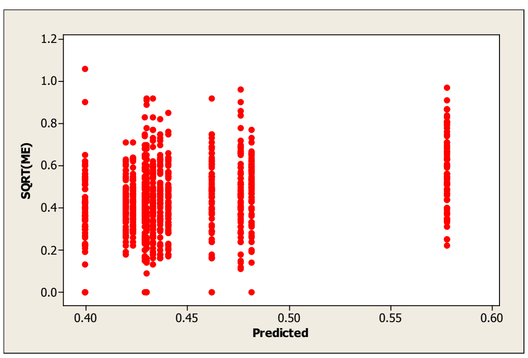
G.5.2.1 Test of Normality



G.5.2.2 Test of Homogeneity of Variance



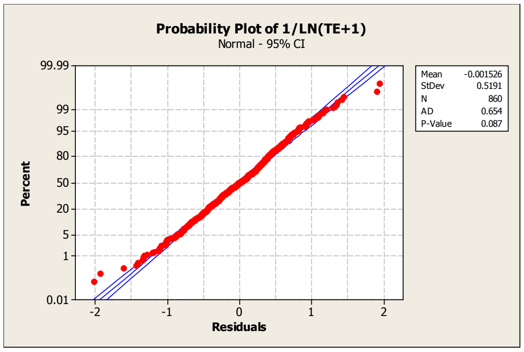
G.5.2.3 Goodness of Fit



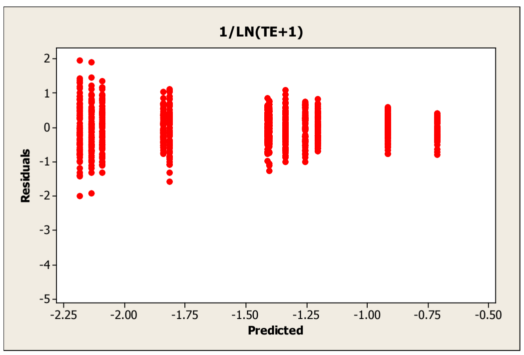
G.6 Typematic Events

G.6.1 TE Test of Assumptions for Final Model

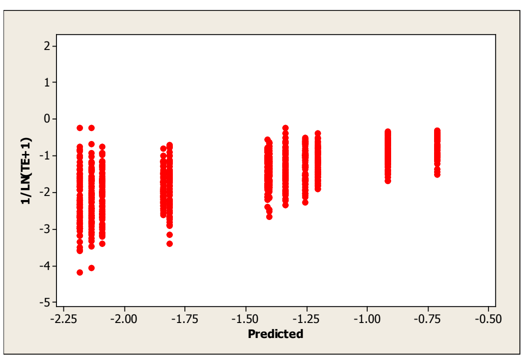
G.6.1.1 Test of Normality



G.6.1.2 Test of Homogeneity of Variance



G.6.1.3 Goodness of Fit



G.7 Keyboard Rank

Friedman Test: Rank versus Keyboard Blocked by Participant

S = 15.08 DF = 2 P = 0.001

Keyboard	N	Est Median	Ranks
Alphabetical	24	2.0000	52.0
Predictive	24	2.6667	59.0
Hybrid	24	1.3333	33.0

Grand median = 2.0000

Rank averages and totals

Level	Rank average	Rank total
1	2.17	52
2	2.46	59
3	1.38	33

Friedman's S = 362

CHI(2)= 15.08, p=0.0005 (No correction for ties.)
No ties in this data set.

==

Table of rank frequencies

Rank	Conditions 1	2	3
1.0	5	3	16
1.5	0	0	0
2.0	10	7	7
2.5	0	0	0
3.0	9	14	1
Rank Tot	52	59	33
Rank Ave	2.2	2.5	1.4

==

Multiple Comparisons Based on Friedman Rank-Sums

Condition A	Condition B	Absolute Difference	Probability Level
1 : 52	2 : 59	7	>.20
1 : 52	3 : 33	19	<.025
2 : 59	3 : 33	26	<.001

G.8 Rank vs. Rating

Correlations: ABCRank, ABCRating

Pearson correlation of ABCRank and ABCRating = 0.050
P-Value = 0.817

Correlations: FOCLRank, FOCLRating

Pearson correlation of FOCLRank and FOCLRating = 0.483
P-Value = 0.017

Correlations: HYBRank, HYBRating

Pearson correlation of HYBRank and HYBRating = 0.297
P-Value = 0.159

Appendix H. Expert Performance Results[8]

H.1 CWPM

H.1.1 CWPM Power Law Model using Session

H.1.1.1 Mixed Effects Model Results

Linear mixed-effects model fit by REML
Data: Exp.3.copyV2.df

AIC	BIC	logLik
-3669.243	-3606.807	1845.621

Random effects:
Formula: ~ log(Session) | Participant
Structure: General positive-definite

	StdDev	Corr
(Intercept)	0.23768244	(Intr)
log(Session)	0.04066696	-0.601
Residual	0.10881820	

Correlation Structure: AR(1)
Formula: ~ 1 | Participant
Parameter estimate(s):
Phi
0.1717939

Variance function:
Structure: Different standard deviations per stratum
Formula: ~ 1 | Keyboard
Parameter estimates:

FOCL	ABC	Hybrid
1	0.8497441	0.9528714

Fixed effects: log(CWPM) ~ log(Session) + Keyboard.ff

	Value	Std.Error	DF	t-value	p-value
(Intercept)	1.548166	0.06906347	2145	22.41657	<.0001
log(Session)	0.166114	0.01217874	2145	13.63969	<.0001
Keyboard.ffFOCL	0.115059	0.00580549	2145	19.81905	<.0001
Keyboard.ffHybrid	0.120458	0.00564714	2145	21.33082	<.0001

Standardized Within-Group Residuals:

Min	Q1	Med	Q3	Max
-10.97209	-0.6115564	0.03277771	0.6387455	3.873501

Number of Observations: 2160
Number of Groups: 12

Analysis of Variance Table

	numDF	denDF	F-value	p-value
(Intercept)	1	2145	1574.393	<.0001
log(Session)	1	2145	186.032	<.0001
Keyboard.ff	2	2145	296.577	<.0001

8 ABC= Alphabetical, FOCL= Predictive, HYB= Hybrid

Approximate 95% confidence intervals

Fixed effects:

	lower	est.	upper
(Intercept)	1.4127281	1.5481664	1.6836048
log(Session)	0.1422309	0.1661142	0.1899976
Keyboard.ffFOCL	0.1036744	0.1150594	0.1264443
Keyboard.ffHybrid	0.1093838	0.1204582	0.1315327

Random Effects:
Level: Participant

	lower	est.	upper
sd((Intercept))	0.15519022	0.23768244	0.36402386
sd(log(Session))	0.02586755	0.04066696	0.06393344
cor((Intercept),log(Session))	-0.86791418	-0.60087487	-0.06437850

Correlation structure:

	lower	est.	upper
Phi	0.1269467	0.1717939	0.21594

Variance function:

	lower	est.	upper
ABC	0.7841950	0.8497441	0.9207723
Hybrid	0.8780045	0.9528714	1.0341221

Within-group standard error:

lower	est.	upper
0.1028736	0.1088182	0.1151063

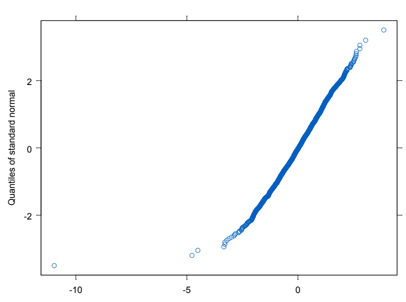
anova(CWPM.PowerLaw, L=c(Keyboard.ffFOCL=1, Keyboard.ffHybrid=-1))
F-test for linear combination(s)

Keyboard.ffFOCL	Keyboard.ffHybrid
1	-1

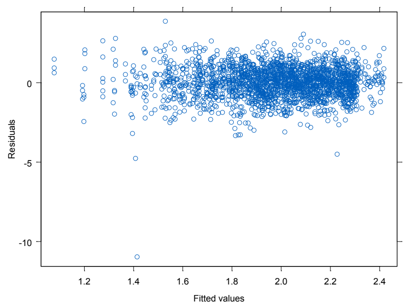
	numDF	denDF	F-value	p-value
1	1	2145	0.7824312	0.3765

H.1.1.2 CWPM Power Law Model Assumptions

H.1.1.2.1 Test of Normality



H.1.1.2.2 Test of Homogeneity of Variance



H.1.2 CWPM Expanded Power Law Model using Session

H.1.2.1 Linear Mixed Effects Model Results

Linear mixed-effects model fit by maximum likelihood
Data: Exp.3.copyV2.df
KeyboardF= FOCL/Predictive, KeyboardH= HYB/Hybrid

AIC	BIC	logLik
-3757.276	-3655.075	1896.638

Random effects:
Formula: ~ log(Session) + Keyboard | Participant
Structure: General positive-definite

	StdDev	Corr		
(Intercept)	0.22231042	(Intr)	lg(Ss)	Kybrd1
log(Session)	0.03952037	-0.633		
Keyboard1	0.05372439	-0.012	0.252	
Keyboard2	0.04776794	0.233	0.039	0.967
Residual	0.09843374			

Correlation Structure: AR(1)
Formula: ~ 1 | Participant
Parameter estimate(s):
Phi
0.1358394

Fixed effects: log(CWPM) ~ log(Session) * Keyboard

	Value	Std.Error	DF	t-value	p-value	
(Intercept)	1.565237	0.06528413	2143	23.97577	<.0001	
log(Session)	0.158178	0.01250476	2143	12.64941	<.0001	
KeyboardF	0.097562	0.02223058	2143	4.38865	<.0001	FOCL-ABC
KeyboardH	0.083796	0.02106411	2143	3.97814	0.0001	Hybrid-ABC
log(Session)KeyboardF	0.008011	0.00703699	2143	1.13835	0.2551	FOCL-ABC
log(Session)KeyboardH	0.017258	0.00703773	2143	2.45220	0.0143	Hybrid-ABC

Standardized Within-Group Residuals:

Min	Q1	Med	Q3	Max
-10.34491	-0.5877723	0.03869423	0.6096603	3.838847

Number of Observations: 2160
Number of Groups: 12

Analysis of Variance Table

	numDF	denDF	F-value	p-value

(Intercept)	1	2143	1707.617	<.0001
log(Session)	1	2143	260.598	<.0001
Keyboard	2	2143	33.056	<.0001
log(Session):Keyboard	2	2143	3.012	0.0494

Approximate 95% confidence intervals

Fixed effects:

	lower	est.	upper
(Intercept)	1.437388233	1.565237137	1.69308604
log(Session)	0.133689203	0.158177845	0.18266649
KeyboardF	0.054027159	0.097562333	0.14109751
KeyboardH	0.042545123	0.083795930	0.12504674
log(Session)KeyboardF	-0.005770311	0.008010558	0.02179143
log(Session)KeyboardH	0.003475603	0.017257907	0.03104021

Random Effects:
Level: Participant

	lower	est.	upper
sd((Intercept))	0.14972274	0.22231042	0.33008961
sd(log(Session))	0.02599325	0.03952037	0.06008713
sd(KeyboardF)	0.03493561	0.05372439	0.08261799
sd(KeyboardH)	0.03071679	0.04776794	0.07428432
cor((Intercept),log(Session))	-0.85551964	-0.63309426	-0.21340736
cor((Intercept),KeyboardF)	-0.17865463	-0.01213074	0.15506875
cor((Intercept),KeyboardH)	-0.07222240	0.23307523	0.49843840
cor(log(Session),KeyboardF)	-0.04968810	0.25234803	0.51209620
cor(log(Session),KeyboardH)	-0.09822086	0.03909137	0.17494294
cor(KeyboardF,KeyboardH)	0.77758999	0.96709210	0.99553637

Correlation structure:

	lower	est.	upper
Phi	0.09084943	0.1358394	0.180276

Within-group standard error:

lower	est.	upper
0.09543785	0.09843374	0.1015237

anova(CWPM.PowerLaw, L=c(KeyboardF=1, KeyboardH=-1))
F-test for linear combination(s)
KeyboardF KeyboardH
1 -1

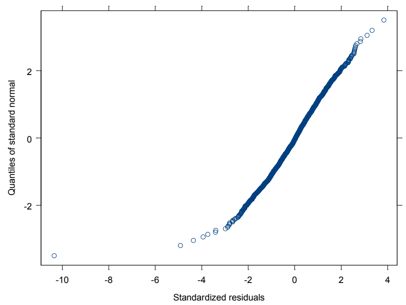
	numDF	denDF	F-value	p-value
1	1	2143	0.7017291	0.4023

> anova(CWPM.PowerLaw, L=c("log(Session)KeyboardF"=1, "log(Session)KeyboardH"=-1))
F-test for linear combination(s)
log(Session)KeyboardF log(Session)KeyboardH
1 -1

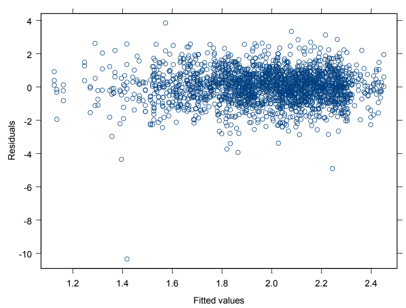
	numDF	denDF	F-value	p-value
1	1	2143	1.727578	0.1889

H.1.2.2 CWPM Expanded Power Law Model Assumptions

H.1.2.2.1Test of Normality



H.1.2.2.2 Test of Homogeneity of Variance



H.1.3 CWPM Quadratic Model using Session

H.1.3.1 Linear Mixed Effects Model Results

```
Linear mixed-effects model fit by REML
 Data: Exp.3.copyV2
       AIC          BIC          logLik
  4781.718     4900.885     -2369.859

Random effects:
 Formula:  ~ Session + Keyboard.f | Participant
 Structure: General positive-definite
                    StdDev           Corr
     (Intercept)    1.13307941       (Intr) Sessin KyFOCL
         Session    0.03468365       0.078
   Keyboard.fFOCL   0.43775520       0.222  0.604
 Keyboard.fHybrid   0.40346897       0.441  0.565  0.973
        Residual    0.82515534

Correlation Structure: AR(1)
 Formula:  ~ 1 | Participant
 Parameter estimate(s):
      Phi
 0.150786
```

Variance function:
Structure: Different standard deviations per stratum
Formula: ~ 1 | Keyboard.f
Parameter estimates:

FOCL	ABC	Hybrid
1	0.6477405	0.9647933

Fixed effects: CWPM ~ Session + Session^2 + Keyboard.f * Session

	Value	Std.Error	DF	t-value	p-value	
(Intercept)	4.876023	0.3331633	2142	14.63554	<.0001	
Session	0.314311	0.0160301	2142	19.60758	<.0001	
I(Session^2)	-0.009125	0.0005676	2142	-16.07659	<.0001	
Keyboard.fFOCL	0.565791	0.1508958	2142	3.74955	0.0002	FOCL-ABC
Keyboard.fHybrid	0.547498	0.1415294	2142	3.86844	0.0001	HYBR-ABC
SessionKeyboard.fFOCL	0.029514	0.0068811	2142	4.28923	<.0001	FOCL-ABC
SessionKeyboard.fHybrid	0.036262	0.0067109	2142	5.40340	<.0001	FOCL-ABC

Standardized Within-Group Residuals:

Min	Q1	Med	Q3	Max
-5.167301	-0.6239767	-0.01170157	0.6363135	4.012337

Number of Observations: 2160
Number of Groups: 12

Analysis of Variance Table

	numDF	denDF	F-value	p-value
(Intercept)	1	2142	303.6453	<.0001
Session	1	2142	128.1115	<.0001
I(Session^2)	1	2142	258.4116	<.0001
Keyboard.f	2	2142	28.7993	<.0001
Keyboard.f:Session	2	2142	18.4288	<.0001

Approximate 95% confidence intervals

Fixed effects:

	lower	est.	upper
(Intercept)	4.22266615	4.876023440	5.529380731
Session	0.28287454	0.314310631	0.345746724
I(Session^2)	-0.01023780	-0.009124737	-0.008011673
Keyboard.fFOCL	0.26987334	0.565790983	0.861708623
Keyboard.fHybrid	0.26994833	0.547497639	0.825046949
SessionKeyboard.fFOCL	0.01602022	0.029514497	0.043008775
SessionKeyboard.fHybrid	0.02310105	0.036261564	0.049422081

Random Effects:
Level: Participant

	lower	est.	upper
sd((Intercept))	0.70319847	1.13307941	1.82575618
sd(Session)	0.02193084	0.03468365	0.05485223
sd(Keyboard.fFOCL)	0.27228370	0.43775520	0.70378658
sd(Keyboard.fHybrid)	0.24806630	0.40346897	0.65622461
cor((Intercept),Session)	-0.60069441	0.07818317	0.69154929
cor((Intercept),Keyboard.fFOCL)	-0.38259462	0.22229244	0.69379112
cor((Intercept),Keyboard.fHybrid)	-0.16570790	0.44094298	0.80548993
cor(Session,Keyboard.fFOCL)	-0.07510157	0.60376969	0.90021644
cor(Session,Keyboard.fHybrid)	-0.17126332	0.56546794	0.89660621
cor(Keyboard.fFOCL,Keyboard.fHybrid)	0.76969209	0.97283238	0.99708989

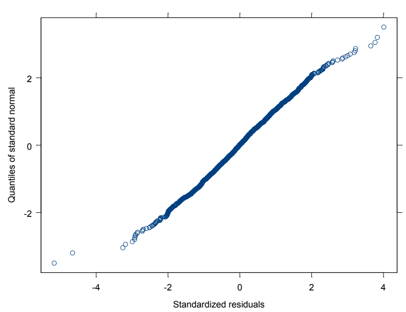
Correlation structure:

	lower	est.	upper
Phi	0.1070742	0.150786	0.1939161

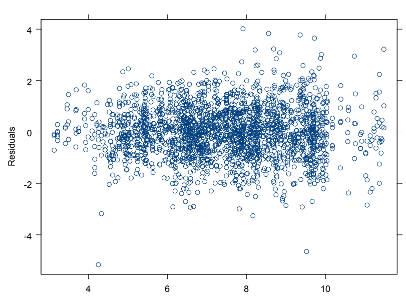
```
Variance function:
          lower        est.         upper
   ABC    0.6004632    0.6477405    0.6987402
Hybrid    0.8897067    0.9647933    1.0462169

 Within-group standard error:
    lower        est.         upper
 0.7808791    0.8251553    0.871942

anova(CWPM.trt.cont, L=c(Keyboard.fFOCL=1, Keyboard.fHybrid=-1))
F-test for linear combination(s)
 Keyboard.fFOCL Keyboard.fHybrid
            1             -1
  numDF denDF     F-value          p-value
1     1    2142   0.03311423       0.8556


anova(CWPM.trt.cont, L=c(SessionKeyboard.fFOCL=1, SessionKeyboard.fHybrid=-1))
F-test for linear combination(s)
 SessionKeyboard.fFOCL SessionKeyboard.fHybrid
                  1                      -1
  numDF denDF   F-value     p-value
1       1   2142    0.7100845   0.3995
```

H.1.3.2 CWPM Quadratic Model Assumptions

H.1.3.2.1 Test of Normality



H.1.3.2.2 Test of Homogeneity of Variance



H.1.4 Quadratic Model for CWPM using Actual Time

H.1.4.1 Linear Mixed Effects Model Results

Linear mixed-effects model fit by REML
Data: Exp.3.copyV2.df

AIC	BIC	logLik
4928.111	5047.277	-2443.055

Random effects:
Formula: ~ CumTime.Days + Keyboard.ff | Participant
Structure: General positive-definite

	StdDev	Corr		
(Intercept)	1.13751000	(Intr)	CmTm.D	KyFOCL
CumTime.Days	0.03587018	0.086		
Keyboard.ffFOCL	0.44610685	0.229	0.585	
Keyboard.ffHybrid	0.42016708	0.425	0.504	0.976
Residual	0.86516988			

Correlation Structure: AR(1)
Formula: ~ 1 | Participant
Parameter estimate(s):
Phi
0.2296426
Variance function:
Structure: Different standard deviations per stratum
Formula: ~ 1 | Keyboard.ff
Parameter estimates:

FOCL	ABC	Hybrid
1	0.6547346	0.9634104

Fixed effects: CWPM ~ Keyboard.ff + CumTime.Days + CumTime.Days^2 + Keyboard.ff * CumTime.Days

	Value	Std.Error	DF	t-value	p-value
(Intercept)	5.395594	0.3329502	2142	16.20541	<.0001
Keyboard.ffFOCL	0.629198	0.1506425	2142	4.17676	<.0001
Keyboard.ffHybrid	0.612894	0.1428778	2142	4.28964	<.0001
CumTime.Days	0.226239	0.0147604	2142	15.32744	<.0001
I(CumTime.Days^2)	-0.005125	0.0004872	2142	-10.51902	<.0001
Keyboard.ffFOCLCumTime.Days	0.025383	0.0065188	2142	3.89378	0.0001
Keyboard.ffHybridCumTime.Days	0.032037	0.0062741	2142	5.10624	<.0001

Standardized Within-Group Residuals:

Min	Q1	Med	Q3	Max
-5.427755	-0.6219275	-0.008422852	0.6309702	4.084469

Number of Observations: 2160
Number of Groups: 12

Analysis of Variance Table

	numDF	denDF	F-value	p-value
(Intercept)	1	2142	281.6934	<.0001
Keyboard.ff	2	2142	8.1924	0.0003
CumTime.Days	1	2142	158.8328	<.0001
I(CumTime.Days^2)	1	2142	111.7666	<.0001
Keyboard.ff:CumTime.Days	2	2142	15.2024	<.0001

Approximate 95% confidence intervals

Fixed effects:

	lower	est.	upper
(Intercept)	4.742655194	5.395594456	6.048533718
Keyboard.ffFOCL	0.333777507	0.629198334	0.924619160
Keyboard.ffHybrid	0.332700616	0.612894247	0.893087878
CumTime.Days	0.197292785	0.226238961	0.255185137
I(CumTime.Days^2)	-0.006080367	-0.005124922	-0.004169478
Keyboard.ffFOCLCumTime.Days	0.012598956	0.025382829	0.038166702
Keyboard.ffHybridCumTime.Days	0.019733235	0.032037271	0.044341308

Random Effects:
Level: Participant

	lower	est.	upper
sd((Intercept))	0.75515408	1.13751000	1.7134636
sd(CumTime.Days)	0.02215486	0.03587018	0.0580762
sd(Keyboard.ffFOCL)	0.27599308	0.44610685	0.7210736
sd(Keyboard.ffHybrid)	0.25673030	0.42016708	0.6876492
cor((Intercept),CumTime.Days)	-0.42363956	0.08633059	0.5547461
cor((Intercept),Keyboard.ffFOCL)	-0.38226549	0.22918918	0.7010576
cor((Intercept),Keyboard.ffHybrid)	-0.17288325	0.42538537	0.7943588
cor(CumTime.Days,Keyboard.ffFOCL)	-0.03616233	0.58471331	0.8799098
cor(CumTime.Days,Keyboard.ffHybrid)	-0.15823411	0.50414828	0.8536024
cor(Keyboard.ffFOCL,Keyboard.ffHybrid)	0.74690880	0.97579274	0.9979299

Correlation structure:

	lower	est.	upper
Phi	0.1862046	0.2296426	0.2721846

Variance function:

	lower	est.	upper
ABC	0.6084018	0.6547346	0.7045959
Hybrid	0.8960766	0.9634104	1.0358038

Within-group standard error:

lower	est.	upper
0.8206603	0.8651699	0.9120935

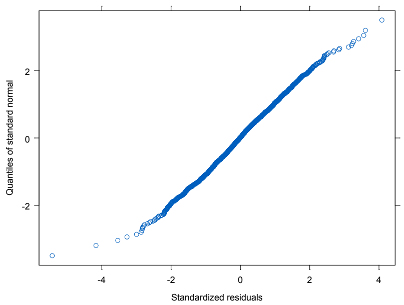
anova(CWPM.CumTime, L=c(Keyboard.ffFOCLCumTime.Days=1, Keyboard.ffHybridCumTime.Days=-1))
F-test for linear combination(s)

Keyboard.ffFOCLCumTime.Days	Keyboard.ffHybridCumTime.Days
1	-1

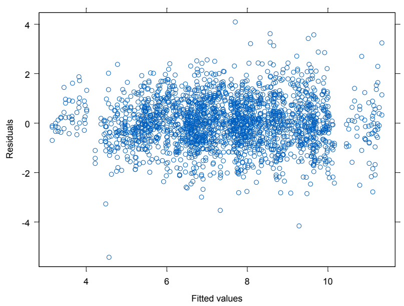
numDF	denDF	F-value	p-value
1	2142	0.8816126	0.3479

H.1.4.2 CWPM Quadratic (using Time) Model Assumptions

H.1.4.2.1 Test of Normality



H.1.4.2.2 Test of Homogeneity of Variance



H.1.5 Quadratic Model for Reaction Time (RT) using Session

H.1.5.1 Linear Mixed Effects Model Results

Data: RTData

AIC	BIC	logLik
207.4748	343.565	-79.73738

Random effects:
Formula: ~ Session + Session^2 + Keyboard | Participant
Structure: General positive-definite

	StdDev	Corr			
(Intercept)	0.739626561	(Intr)	Session	I(S^2)	Kybrd1
Session	0.066827351	-0.463			
I(Session^2)	0.002208224	0.452	-0.995		
Keyboard1	0.258709550	-0.145	0.496	-0.557	
Keyboard2	0.215388228	-0.483	0.676	-0.729	0.854
Residual	0.255005546				

Correlation Structure: AR(1)
Formula: ~ 1 | Participant
Parameter estimate(s):
Phi
0.234558

Variance function:
Structure: Different standard deviations per stratum
Formula: ~ 1 | Keyboard
Parameter estimates:

FOCL	ABC	Hybrid
1	1.104485	0.8129796

Fixed effects: 1/RT ~ Session + Session^2 + Keyboard

	Value	Std.Error	DF	t-value	p-value	
(Intercept)	3.795788	0.2148672	2133	17.66574	<.0001	
Session	0.157379	0.0198750	2133	7.91845	<.0001	
I(Session^2)	-0.004372	0.0006749	2133	-6.47792	<.0001	
KeyboardF	-0.816930	0.0763350	2133	-10.70190	<.0001	FOCL-ABC
KeyboardH	-1.243045	0.0638589	2133	-19.46547	<.0001	Hybrid-ABC

Standardized Within-Group Residuals:

Min	Q1	Med	Q3	Max
-8.125042	-0.5974939	0.02638843	0.6202213	4.625512

Number of Observations: 2149
Number of Groups: 12

Analysis of Variance Table

	numDF	denDF	F-value	p-value
(Intercept)	1	2133	108.1769	<.0001
Session	1	2133	41.0707	<.0001
I(Session^2)	1	2133	774.0309	<.0001
Keyboard	2	2133	244.6619	<.0001

Approximate 95% confidence intervals

Fixed effects:

	lower	est.	upper
(Intercept)	3.374416817	3.795787965	4.217159114
Session	0.118402906	0.157379339	0.196355772
I(Session^2)	-0.005695201	-0.004371735	-0.003048269
KeyboardF	-0.966628378	-0.816929580	-0.667230783
KeyboardH	-1.368276995	-1.243044697	-1.117812400

Random Effects:
Level: Participant

	lower	est.	upper
sd((Intercept))	0.483752165	0.739626561	1.13084238
sd(Session)	0.042061444	0.066827351	0.10617550
sd(I(Session^2))	0.001353076	0.002208224	0.00360383
sd(Keyboard1)	0.166316173	0.258709550	0.40243008
sd(Keyboard2)	0.131688451	0.215388228	0.35228669
cor((Intercept),Session)	-0.838976458	-0.462627167	0.21309327
cor((Intercept),I(Session^2))	-0.259920539	0.452101589	0.84565525
cor((Intercept),Keyboard1)	-0.766219849	-0.145496827	0.61569025
cor((Intercept),Keyboard2)	-0.865596675	-0.482616421	0.25660099
cor(Session,I(Session^2))	-0.999535363	-0.995195521	-0.95130745
cor(Session,Keyboard1)	-0.134478105	0.496361161	0.84089746
cor(Session,Keyboard2)	0.133658669	0.676388768	0.90700843
cor(I(Session^2),Keyboard1)	-0.870490633	-0.557443882	0.07671644
cor(I(Session^2),Keyboard2)	-0.930870717	-0.729114456	-0.18658279
cor(Keyboard1,Keyboard2)	0.495113551	0.853748595	0.96380265

Correlation structure:

	lower	est.	upper
Phi	0.189362	0.234558	0.2787622

Variance function:

	lower	est.	upper
ABC	1.0243591	1.1044847	1.1908777
Hybrid	0.7546582	0.8129796	0.8758082

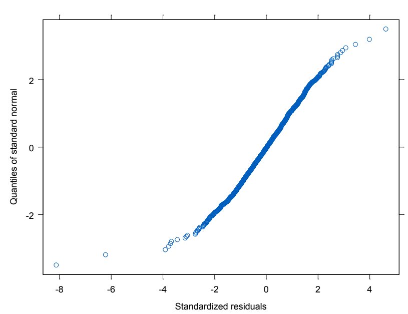
Within-group standard error:

lower	est.	upper
0.2414529	0.2550055	0.269319

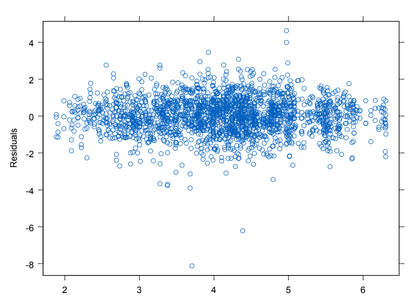
```
anova(RT.lme, L=c(Keyboard1=1, Keyboard2=-1))
F-test for linear combination(s)
 Keyboard1 Keyboard2
        1        -1
    numDF denDF F-value    p-value
 1     1      2133     106.729  <.0001
```

H.1.5.2 Reaction Time Model Assumptions

H.1.5.2.1 Test of Normality



H.1.5.2.2 Test of Homogeneity of Variance



H.2 KSPC

H.2.1 Linear Mixed Effects Model Results

Linear mixed-effects model fit by REML
Data: Exp.3.copyV2

AIC	BIC	logLik
4432.836	4563.202	-2193.418

Random effects:
Formula: ~ Session + Keyboard | Participant
Structure: General positive-definite

	StdDev	Corr		
(Intercept)	0.95740319	(Intr)	Sessin	KyFOCL
Session	0.01738942	-0.609		
KeyboardFOCL	0.70344719	-0.974	0.502	
KeyboardHybrid	0.56175566	-0.932	0.571	0.978
Residual	0.70113040			

Variance function:
Structure: Different standard deviations per stratum
Formula: ~ 1 | Keyboard
Parameter estimates:

FOCL	ABC	Hybrid
1	0.9578376	0.8156456

Fixed effects: KSPC ~ Session + SessionSQR + Keyboard * Session + Keyboard * SessionSQR + Gender

	Value	Std.Error	DF	t-value	p-value
(Intercept)	9.234285	0.2938205	2129	31.42832	<.0001
Session	-0.153137	0.0189748	2129	-8.07055	<.0001
SessionSQR	0.005231	0.0008481	2129	6.16863	<.0001
KeyboardFOCL	-2.834092	0.2361100	2129	-12.00327	<.0001
KeyboardHybrid	-3.680600	0.1956412	2129	-18.81301	<.0001
Gender	0.254239	0.1095951	10	2.31980	0.0428
SessionKeyboardFOCL	0.109459	0.0264527	2129	4.13790	<.0001
SessionKeyboardHybrid	0.139765	0.0240395	2129	5.81396	<.0001
SessionSQRKeyboardFOCL	-0.004097	0.0012259	2129	-3.34212	0.0008
SessionSQRKeyboardHybrid	-0.005280	0.0011141	2129	-4.73879	<.0001

Standardized Within-Group Residuals:

Min	Q1	Med	Q3	Max
-4.138933	-0.6373405	-0.02376274	0.5868277	5.339122

Number of Observations: 2149
Number of Groups: 12

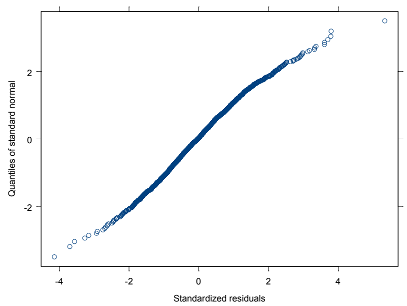
Analysis of Variance Table

	numDF	denDF	F-value	p-value
(Intercept)	1	2129	13917.90	<.0001
Session	1	2129	88.54	<.0001
SessionSQR	1	2129	16.15	0.0001
Keyboard	2	2129	423.26	<.0001
Gender	1	10	5.38	0.0429
Keyboard:Session	2	2129	13.71	<.0001
Keyboard:SessionSQR	2	2129	11.72	<.0001

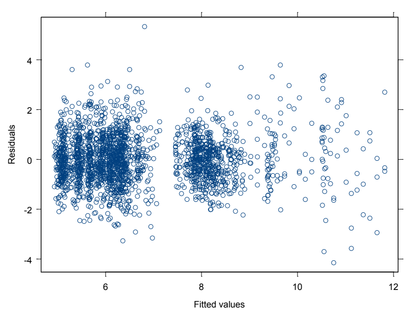
```
Approximate 95% confidence intervals

 Fixed effects:
                                   lower          est.           upper
           (Intercept)             8.658079647    9.234284848    9.810490048
             Session               -0.190347988   -0.153136938   -0.115925889
           SessionSQR              0.003568299    0.005231432    0.006894564
          KeyboardFOCL             -3.297122288   -2.834092044   -2.371061800
         KeyboardHybrid            -4.064267805   -3.680599939   -3.296932072
             Gender                0.010045613    0.254238773    0.498431933
     SessionKeyboardFOCL           0.057582866    0.109458676    0.161334486
    SessionKeyboardHybrid          0.092621432    0.139764805    0.186908178
   SessionSQRKeyboardFOCL -0.006501439            -0.004097260   -0.001693082
SessionSQRKeyboardHybrid -0.007464525             -0.005279629   -0.003094733

 Random Effects:
  Level: Participant
                                          lower          est.           upper
             sd((Intercept))              0.397685429    0.95740319     2.30488922
                sd(Session)               0.006787013    0.01738942     0.04455451
            sd(KeyboardFOCL)              0.298183291    0.70344719     1.65950932
           sd(KeyboardHybrid)             0.260426382    0.56175566     1.21174135
        cor((Intercept),Session)          -0.939597640   -0.60946776    0.30810972
    cor((Intercept),KeyboardFOCL)         -0.996206224   -0.97370755    -0.82919867
  cor((Intercept),KeyboardHybrid)         -0.998980736   -0.93194548    0.41751823
     cor(Session,KeyboardFOCL)            -0.552032986   0.50172818     0.93840581
    cor(Session,KeyboardHybrid)           -0.223610520   0.57137166     0.90983535
cor(KeyboardFOCL,KeyboardHybrid)  0.798203726            0.97809459     0.99781682

 Variance function:
         lower          est.           upper
  ABC    0.8803353      0.9578376      1.0421630
Hybrid   0.7540131      0.8156456      0.8823159

 Within-group standard error:
    lower          est.           upper
 0.6628399      0.7011304      0.7416329


F-test for linear combination(s)
 KeyboardFOCL KeyboardHybrid
        1       -1
     numDF denDF  F-value    p-value
1    1      2129  45.58277   <.0001

anova(KSPC.trt.cont, L=c(SessionKeyboardFOCL=1, SessionKeyboardHybrid=-1))

F-test for linear combination(s)
 SessionKeyboardFOCL SessionKeyboardHybrid
                1                -1
   numDF denDF  F-value    p-value
 1    1      2129  1.510687  0.2192


anova(KSPC.trt.cont, L=c(SessionSQRKeyboardFOCL=1, SessionSQRKeyboardHybrid=-1))

F-test for linear combination(s)
 SessionSQRKeyboardFOCL SessionSQRKeyboardHybrid
                  1                 -1
              numDF denDF  F-value p-value
   1          1       2129  1.070549  0.3009
```

H.2.2 KSPC Test of Model Assumptions

H.2.2.1 Test of Normality



H.2.2.2 Test of Homogeneity of Variance



H.3 Movement Inefficiency

H.3.1 Computed Minimal Path

Phrase	Alphabetical	Predictive	Hybrid
It was late, we were playing kissing games, and Jessica and I were called on to kiss in front of the others.	777	490	434
He declared the government is thinking of asking for foreign troops if the situation worsens.	746	471	366
However, there are still several types of calls that necessitate the use of telephone operators.	703	464	413
Investors studying the toll-road bonds for opportunities find that not all roads are nearing their goals.	830	532	479
For the year to date, sales of the company's farm equipment dealers still lag about 5% behind 1960.	700	504	441
There is so far no evidence to indicate conclusively that this coupling is under enzymatic control.	738	471	430
When Huff attempted to cash another $100 check there Monday, hotel officials called police.	653	466	458
As soon as you find out if they are Geely and Harris, come on around to the lounge where I'll be waiting.	755	442	418
Lucy drew out the chair and sat down; she relaxed a little, and some of the tension went out of her.	694	513	441
Both men knew it was in the Norberg family holdings, but to which of the cousins did it belong, Anta or Freya?	808	540	488
Matsuo had faked, both the burned and the unburned, the latter decomposing rapidly under the tropical sun.	771	529	468
While I respect his sincere concern for peace, he made four points that I would like to question.	735	487	413
A body of water is usually the center of interest at parks which attract the greatest picnic and camping use.	721	539	464
There was evident delight on the part of the subject in response to her experience of the freedom of movement.	808	557	462
By political, economic, geographic and natural standards, they were justified in doing so.	683	508	416
At one astronomical unit from the sun (the Earth's distance) the dust orbits are probably nearly circular.	763	553	514
Russia's young gymnasts have studied dance before having the rigorous training on apparatus.	678	531	467
But this was not unusual, because youth in these quarters was always pushed at a distance from its elders.	701	504	477
Johnny vigorously pounded two bleached steer bones against the gourd which served as his drum.	678	448	411
It is proposed that in the future complete sampling censuses be carried out at five-year intervals.	720	491	468
In most Western cultures today these twins have been sent away to the libraries and museums.	642	442	452
Apart from journalese and vaudeville gags, the anatomical is also found in jocular literature.	704	470	419
Among the spectators was the noted exotic dancer, Patti Waggin who is Mrs. Don Rudolph when off the stage.	780	593	524
Their appearance, next spring, coincides in an almost uncanny way with the flowering of their host plants.	810	568	478

Phrase	Alphabetical	Predictive	Hybrid
The stress on have, which here represents have finished reading the paper, is quite strong.	675	417	375
Chandler left Carroll at the bottom of the hill to direct any reinforcements he could find to the fight.	770	495	424
There was a very old man and a young woman and a brood of children ranging from toddlers to teen-agers.	730	451	408
But don't tell that to a veteran of the Fighting Seventh, especially in a saloon on Saturday night.	732	511	483
It squatted low and square upon the sidewalk with a heavy iron grating supporting a glass facade.	672	496	433
The ordinance would increase fees from $1 for males and $2 for females to a flat $5 a dog.	611	408	389
But remember this - it isn't the aircraft which is vulnerable to nuclear rockets, it is the airfield.	727	549	452
Similar payroll tax boosts would be imposed on those under the railroad retirement system.	695	476	450
If we return to them today, we have no difficulty spotting their weaknesses but we find them still pleasing.	786	564	492
The highroad, one might say at first, belongs to life, while the way to the churchyard belongs to death.	737	505	462
In the average situation about one-third of those visited make commitments to Christ and the Church.	745	490	453
I have done everything, he wrote, to break up the whole of that unfortunate establishment.	661	474	418
The pace could now be accelerated, for the inhabitants of the Aegean stood on firm ground.	638	460	383
To my immense relief, she changed the subject in the next sentence: shall we go to the Louvre today?	723	493	450
He pulled it over, climbed up, and lifted out the big volume, almost losing his balance from the weight of it.	825	600	480
It was the best he could hope for on a watch that had ended with a session in Killpath's office.	660	435	380
The Cunard line has under consideration replacing the Queen Mary with a ship smaller than 75,000 tons.	759	480	476
Winston took out a pencil, admired the point, and wrote slowly and heavily, clothes stand.	700	448	370
Be sure that the landing foot is brought close to the hands and that only one foot lands at a time.	696	401	350
Not necessarily to be off all by ourselves, but away from the crowds and common happenstance.	669	497	457
A black, snake-like object swayed eerily in front of him, spewing bubbles from its flat cobra head.	704	595	497
No amount of religious ceremonies or even joining a church will relieve the gnawing of your inner space.	817	587	457
This, together with a derby hat and horn-rim eyeglasses, gave me the appearance of a Russian nihilist.	728	540	491
So each reading can be given a weight and each reading a score by adding up these weights.	554	443	392
The reality of spirit emerges in this play in spite of the author's convictions to the contrary.	732	492	423
John Heffernan, playing Larry Larkin, the cartoonist, carries the show in marvelous fashion.	750	510	454

Phrase	Alphabetical	Predictive	Hybrid
How to feed: for prevention of ketosis, feed 1/4 pound per head daily for 6 weeks commencing at calving time.	809	573	500
From 1 July 1958 to 30 June 1960, 24 numbers of the Journal and nine of the Bulletin were published.	740	501	474
But he decided he wouldn't mind company in return for free drinks, even though he made good money at his job.	803	573	488
In front of you is the Palazzo Madama, once belonging to the Medici and now the Italian Senate.	701	463	445
Understanding a work of art involves recognition of the ideas that it reflects or embodies.	682	503	397
That is an evening of music-making that would faze many a younger man; Mr. Elman is 70 years old.	702	485	458
The third name was (John) Ravencroft, who was admitted to the Inner Temple in November 1631.	684	465	480
Sports Writer Ensign Ritchie of the Ogden Standard Examiner went to his compartment to talk with him.	789	510	487
It was a real stimulant to a lot of guys I know who have moved past the 2-score-year milestone.	674	471	438
Dear sirs: let me begin by clearing up any possible misconception in your minds, wherever you are.	731	521	479
He is publicly on record as believing Mr. De Sapio should be replaced for the good of the party.	687	555	488
America, America, God shed His grace on thee, and crown thy good with brotherhood from sea to shining sea.	776	525	474
Years later, franks-in-buns were accepted as the first to go at the New York Polo Grounds.	659	463	457
Work that might cost $500 to $750 in the South could cost $750 to $1,200 in New York City or Chicago.	704	517	566
Nevertheless, in another way modern historians still labor in the vineyard of the Oxford school.	752	488	446
The Secretary of the Interior may issue rules and regulations to effectuate the purposes of this Act.	721	566	520
The moonlit night was made for romance, and he had been looking at her soulfully for some time.	733	444	364
Now 38, Mr. Simpkins was graduated from the University of Maryland's College of Agriculture in 1947.	783	602	563
But he knew; he sniffed the air and licked it on his lip and knew as a vintner knows a vintage.	663	454	382
You could think yourself as grown up as Methuselah, yet the maternal voice still kept its comforting magic.	813	564	501
The following discussion of this subject has been adapted from the book Causes Of Catastrophe by L. Don Leet.	792	628	590
Art Lund, a fine big actor with a great head of blond hair and a good voice, impersonates Enright.	689	509	447
The spirit served chiefly to lull the West while Moscow made inroads into the Middle East.	669	449	425
An alternate hatchway entrance, shown on page 25, would reduce the cost of materials $50 to $100.	656	523	513
Information on pages 8 to 14 may help you in deciding on the kind and scale of your farming venture.	745	505	416
A great deal of labor we have as well, for we are too uncertain of where trust may be placed.	589	459	406

Phrase	Alphabetical	Predictive	Hybrid
Yet paradoxically my liberal friends continue to view Jefferson as one of their patron saints.	701	521	480
The window looked out on the Place Redoute - it was the only window of the apartment that did.	672	427	375
The illustration (fig. 11) shows this shelter with the roof at ground level and mounded over.	696	429	402
During the Civil War, Custer, who achieved a brilliant record, was made brigadier general at the age of 23.	716	556	509
I am told that a mortar longer slaked might have remained longer in condition for painting.	718	438	333
We had looked forward to what we hoped to be our first informal meeting with a number of Moscow's artists.	777	530	460
When cutting the pieces, dress the ends smooth, and square with a smooth file or sanding disk.	707	461	394
They hope that if history vouchsafes the West another Budapest, we will receive the opportunity gladly.	734	540	499
For you, readers, are an all-important part of the spiritual experiment that is Guideposts.	701	479	432
So Prokofieff was able to cultivate his musical talents and harvest a rich reward from them.	648	477	421
To the west, the dark green hills of Leyte were lost in the clouds about halfway up their slopes.	693	495	432
He was able to find meaning in his art as long as it was the answer to air raids and gas ovens.	628	403	345
Experts point to the thinning of pitching talent in the American League caused by expansion.	696	476	410
It is most probable that Freud and the Oedipus complex never entered his head in the writing of this story.	773	500	458
The fox is all ingratiating smiles when he arrives from New Orleans, accompanied by one wharf rat.	724	503	460
On Friday he will go to Portland for the swearing in of Dean Bryson as Multnomah County Circuit Judge.	772	562	557
He backed Jess into a corner, grabbed a handful of the man's shirtfront, and drew back his right fist.	715	546	467
He saw that Dolores intended to wait until the last minute, thinking he would get nervous.	683	404	361
After a while he began to feel better about it, especially when no one bothered to ask any questions.	700	486	429
Dams, river development schemes, transportation networks, educational systems require years to construct.	840	661	572
I think I would have been much disappointed in Japan if I had not seen Kyoto, Nara, and Hiroshima.	743	537	482
One day, to everyone's astonishment, someone drops a match in the powder keg and everything blows up.	777	545	479
You can get this added heating feature for as little as $200 more than the price of cooling alone.	672	431	409
The rabbi said thoughtfully, I would not want my people to get in trouble with the Church.	652	463	418
A supplementary grant from the Geological Society of America helped finance its publication.	702	526	458
She seemed so unimpressed that he was obliged to roll up his blue jeans so she could see his brace.	694	475	435

Phrase	Alphabetical	Predictive	Hybrid
I am usually filled with an uneasiness that through some unwitting slip all hell may break loose.	735	458	431
He went swiftly up the sidewalk toward the parked car with the two Beach detectives in the front seat.	668	491	446
They involve only simple mathematics that are taught in grammar school arithmetic classes.	679	478	403
It was mostly for the benefit of the mailman, because hardly anybody else ever visited us.	611	470	419
As faulty as has been our leadership clearly the United States must be relied upon to lead.	614	452	426
I've never done this before, they always said, waiting for the elevator in the hotel corridor.	687	469	417
They remained close together, their air trail wiggling like serpents traveling side by side.	690	466	394
It could, by avoiding direct intervention, provide a short-of-war strategy to meet short-of-war infiltration.	856	633	597
He expected nothing for himself but that which naturally follows those marked for misfortune.	717	458	406
He began to wish that he hadn't shouted that other evening when the truck bore down through the crossing.	755	445	409
It is an irritable rule that does baseball more harm than good, especially at the minor league level.	726	552	433
If this choice is less exciting than New York Democrats may wish, it nevertheless must be made.	683	528	471
I was the first to get my squad on the ball, and anybody thinking it was easy is pretty damn dumb.	650	520	466
Since arriving here, however, I have formed a far different religious picture of present-day England.	752	566	499
The calibration of piezoelectric sensors in terms of the particle parameters is very uncertain.	711	522	443
Dealers would do well to visit such a campground often, look at the equipment and talk with the campers.	768	496	405
Such an understanding, although it must seek to be sympathetic, is not a matter of intuition.	706	470	405
She refolded the letter, replaced it in its envelope, and turned with relief to one from her brother George.	817	549	460
And to offend the dead meant to incur their wrath, and thus provoke the unleashing of countrywide disasters.	782	517	446
I knew the only way I could beat you was to play possum, but it was a good try, kid, and I appreciate it.	694	568	508
When Alec finished reading he was sure that either Forbes or Stacy had killed Diana Beauclerk.	631	445	443
The only evidence of occupation came from the chimney, which was belching out thick smoke.	671	497	404
If we did not mean to say this, why should we be so relieved on finding that the suffering had not occurred?	752	536	474
And with Progressivism the Religion of Humanity was replacing what Gabriel called Christian supernaturalism.	834	616	555
In a lacey open weave shoes have a luster finish, braided collar and bow highlight on the squared throat.	726	524	461
The President was even more generous with the First Lady than he had been before the tragedy.	629	392	392

Phrase	Alphabetical	Predictive	Hybrid
Two men, together like us, we could do something fine out there, maybe find a place where no one's ever been.	808	523	472
He stood looking down at her for a moment, wondering what could have reduced her to this condition.	755	442	381
The other problem is the matter of financing the transition period in the several cities and towns.	757	481	386
A study at the Pentagon and at the service academies revealed that nothing was being done there.	631	423	390
In the darkness he could see the rosy reflection of the neon sign on the wall opposite the window.	746	501	406
The resulting, indescribable torment affects every Southern mind and is the basis of the Southern hysteria.	789	523	470
Flushing stadium in works the New York franchise is headed by Mrs. Charles Shipman Payson.	683	482	492
The man whose reactions and conclusions get the most space is, of course, the Field Marshal himself.	766	539	466
The inadequacy of our library system will become critical unless we act vigorously to correct this condition.	794	580	521
In the same way I like to think we owe our loyalty as legislators to our community, our district, our State.	807	563	486
Having a boat financed through a local bank is done much the same way as an automobile loan is extended.	701	490	438
He had retained his hat and his horn, and, whatever fun might still be going, he was ready to join it.	720	481	403
These machines produce the higher quality stretch yarns required in weaving stretch and textured fabrics.	727	535	505
In the darkness he could see the rosy reflection of the neon sign on the wall opposite the window.	746	501	406
He felt cheerful again, refreshed; presentable in his wide-cut brown suit, the well-made riding boots.	760	574	498
The congressman's patriotism is always involved when he turns upon the Defense Department.	707	493	454
Independent market owners work six days a week; and my husband hasn't had a vacation in 14 years.	653	497	457
He quickly called on Ghana, Tunisia, Morocco, Guinea and Mali, which dispatched troops within hours.	789	543	503
The name alkali bee indicates that one has to look for them in rather inhospitable places.	667	455	355
Our efforts to overcome the lead of the Russians in space are bound to mean accelerated Federal spending.	748	504	459
The unstable political situation there represents one reason new plants shy away from the East Side.	739	507	494
They were staring at him in the same blank and menacing way that the men outside the gate had stared.	661	400	357
Two minutes later it came again- a double explosion, followed by a third, sounding more distant.	728	506	441
Here's a present for you, he said, shoving his bullet-riddled hat down over Nate's purpling forehead.	776	535	496
The sun, blazing hot as prophesied, was far from kind to Mrs. Kirby's varicolored properties.	736	499	467
As critic Walter Kerr points out: adaptations, so long as they are good, still qualify as creative.	726	553	503

Phrase	Alphabetical	Predictive	Hybrid
He is driven back by his yearning to the wintry homeland of his fathers in the forest of Tiveden.	686	445	400
The air took on a special strength now that they'd left the fecund warmth of the farmland behind.	690	505	406
Each time a dictionary form matches a text form, the information cell of the matching text form is saved.	758	544	433
Serve each breast on a thin slice of slow-baked ham and sprinkle with Thompson seedless grapes.	704	493	425
Yet even in these marriages, psychologists say, wives are asserting themselves more strongly.	708	509	465
In the last few years the telephone company has managed to automate many areas of their service.	658	452	413
It should be enough to say that the practice of the state buying automobiles is at least forty years old.	703	493	456
A couple decks for me, Mr. Skyros- and ten-twelve to sell, see, I like to have a little ready cash.	719	544	508
He could not recognize it; he was absolutely unfamiliar with it because he had no visual memory at all.	716	563	486
Goodwin was telegrapher for the American Telegraph Company and the Troy and Canada Junction Telegraph Company.	816	569	574
Are you indiscriminately offering unnecessary medical services- flu shots, sun lamp treatments, etc.?	793	595	528
For example: don't wall in your kitchen before you hang the wall cabinets and set the appliances.	702	460	424
The building will contain 430,000 square feet, approximately the same as our present plant.	670	448	436
That is, they used opaque color throughout, getting solid highlights with active lime white.	705	458	403
In 1872 there were known to be twenty-two in Norton County, and one had been in the family for 200 years.	752	497	515
It was the hard way to fight a war but Thomas did it without making any disastrous mistakes.	611	385	353
In the 1890's the Palace Hotel began serving an oyster dish named after its manager, John C. Kirkpatrick.	785	566	520
It was only the other day that I saw something of yours, about something or other, in some magazine.	742	488	445
In the casual field straws feature wedge heels of cork or carved wood in a variety of styles.	609	522	468
It snowed softly, silently, an undulating interruption of his vision against the night sky.	740	536	422
That keeps in the cold, retains moisture and prevents the heaving of alternate freezing and thawing.	733	477	434
He keeps riding me because I like to listen to the radio and sing while I'm taking a bath.	649	427	358
Within the narrow frame of military tactics, too, the experts agree that the campaign was brilliant.	724	517	434
In Nara I stayed at the hotel where the Prince and Princess had stayed on their honeymoon.	644	430	429
And Lawrence Chase, son of the Ransom Chases, is listed at his new address in Oxford, England.	693	511	505
Of course it was water he really craved; down in the broil of the sun he was becoming dried out.	649	473	396

H.3.2 MI Friedman's Test

H.3.2.1 MI for Session 1

Rank averages and totals

Level	Rank average	Rank total
1: Alphabetical	2.25	27
2: Predictive	2.25	27
3: Hybrid	1.50	18

Friedman's S = 54

CHI(2)= 4.50, p=0.1054 (No correction for ties.)
No ties in this data set.
==
Table of rank frequencies

Rank	Conditions 1	2	3
1.0	3	2	7
1.5	0	0	0
2.0	3	5	4
2.5	0	0	0
3.0	6	5	1
Rank Tot	27	27	18
Rank Ave	2.3	2.3	1.5

==
Multiple Comparisons Based on Friedman Rank-Sums

Condition A	Condition B	Absolute Difference	Probability Level
1 : 27	2 : 27	0	>.20
1 : 27	3 : 18	9	<.20
2 : 27	3 : 18	9	<.20

H.3.2.2 MI Session 10

Rank averages and totals

Level	Rank average	Rank total
1: Alphabetical	1.08	13
2: Predictive	2.50	30
3: Hybrid	2.42	29

Friedman's S = 182
CHI(2)= 15.17, p=0.0005 (No correction for ties.)
No ties in this data set.
==
Table of rank frequencies

Rank	Conditions 1	2	3
1.0	11	0	1
1.5	0	0	0
2.0	1	6	5
2.5	0	0	0
3.0	0	6	6
Rank Tot	13	30	29
Rank Ave	1.1	2.5	2.4

Multiple Comparisons Based on Friedman Rank-Sums

Condition A	Condition B	Absolute Difference	Probability Level
1 : 13	2 : 30	17	<.005
1 : 13	3 : 29	16	<.005
2 : 30	3 : 29	1	>.20

H.3.2.3 MI Session 20

Rank averages and totals

Level	Rank average	Rank total
1: Alphabetical	1.17	14
2: Predictive	2.42	29
3: Hybrid	2.42	29

Friedman's S = 150
CHI(2)= 12.50, p=0.0019 (No correction for ties.)
No ties in this data set.

==

Table of rank frequencies

	Conditions		
Rank	1	2	3
1.0	11	1	0
1.5	0	0	0
2.0	0	5	7
2.5	0	0	0
3.0	1	6	5
Rank Tot	14	29	29
Rank Ave	1.2	2.4	2.4

==

Multiple Comparisons Based on Friedman Rank-Sums

Condition A	Condition B	Absolute Difference	Probability Level
1 : 14	2 : 29	15	<.01
1 : 14	3 : 29	15	<.01
2 : 29	3 : 29	0	>.20

H.3.2.4 Alphabetical MI at Session 1, 10, and 20

Rank averages and totals

Level	Rank average	Rank total
1: Alphabetical	2.83	34
2: Predictive	1.75	21
3 Hybrid	1.42	17

Friedman's S = 158
CHI(2)= 13.17, p=0.0014 (No correction for ties.)
No ties in this data set.

==

Table of rank frequencies

Rank	Conditions 1	2	3
1.0	1	3	8
1.5	0	0	0
2.0	0	9	3
2.5	0	0	0
3.0	11	0	1
Rank Tot	34	21	17
Rank Ave	2.8	1.8	1.4

==

Multiple Comparisons Based on Friedman Rank-Sums

Condition A	Condition B	Absolute Difference	Probability Level
1 : 34	1 : 21	13	<.025
1 : 34	3 : 17	17	<.005
2 : 21	3 : 17	4	>.20

H.3.2.5 Predictive MI at Session 1, 10, and 20

Rank averages and totals

Level	Rank average	Rank total
1: Alphabetical	2.50	30
2: Predictive	1.75	21
3: Hybrid	1.75	21

Friedman's S = 54
CHI(2)= 4.50, p=0.1054 (No correction for ties.)
No ties in this data set.
==

Table of rank frequencies

Rank	Conditions 1	2	3
1.0	2	5	5
1.5	0	0	0
2.0	2	5	5
2.5	0	0	0
3.0	8	2	2
Rank Tot	30	21	21
Rank Ave	2.5	1.8	1.8

==

Multiple Comparisons Based on Friedman Rank-Sums

Condition A	Condition B	Absolute Difference	Probability Level
1 : 30	2 : 21	9	<.20
1 : 30	3 : 21	9	<.20
2 : 21	3 : 21	0	>.20

H.3.2.6 Hybrid MI at Session 1, 10, and 20

Rank averages and totals

Level	Rank average	Rank total
1: Alphabetical	2.00	24
2: Predictive	2.17	26
3: Hybrid	1.83	22

Friedman's S = 8
CHI(2)= 0.67, p=0.7165 (No correction for ties.)
No ties in this data set.
===
Table of rank frequencies

	Conditions		
Rank	1	2	3
1.0	4	4	4
1.5	0	0	0
2.0	4	2	6
2.5	0	0	0
3.0	4	6	2
Rank Tot	24	26	22
Rank Ave	2.0	2.2	1.8

===
Multiple Comparisons Based on Friedman Rank-Sums

Condition A	Condition B	Absolute Difference	Probability Level
1 : 24	2 : 26	2	>.20
1 : 24	3 : 22	2	>.20
2 : 26	3 : 22	4	>.20

H.4 Typematic Events

H.4.1 Friedman's Tests

H.4.1.1 TE at Session 1

Rank averages and totals

Level	Rank average	Rank total
1: Alphabetical	3.00	36
2: Predictive	1.92	23
3: Hybrid	1.08	13

Friedman's S = 266
CHI(2)= 22.17, p=0.0000
(No correction for ties.)
No ties in this data set.

Table of rank frequencies

Rank	Conditions 1	2	3
1.0	0	1	11
1.5	0	0	0
2.0	0	11	1
2.5	0	0	0
3.0	12	0	0
Rank Tot	36	23	13
Rank Ave	3.0	1.9	1.1

==

Multiple Comparisons Based on Friedman Rank-Sums

Condition A	Condition B	Absolute Difference	Probability Level
1 : 36	2 : 23	13	<.025
1 : 36	3 : 13	23	<.0001
2 : 23	3 : 13	10	<.20

H.4.1.2 TE at Session 10

Rank averages and totals

Level	Rank average	Rank total
1: Alphabetical	2.92	35
2: Predictive	1.92	23
3: Hybrid	1.17	14

Friedman's S = 222
CHI(2)= 18.50, p=0.0001 (No correction for ties.)
No ties in this data set.

==

Table of rank frequencies

Rank	Conditions 1	2	3
1.0	0	1	10
1.5	0	0	0
2.0	1	11	2
2.5	0	0	0
3.0	11	0	0
Rank Tot	35	23	14
Rank Ave	2.9	1.9	1.2

==

Multiple Comparisons Based on Friedman Rank-Sums

Condition A	Condition B	Absolute Difference	Probability Level
1 : 35	2 : 23	12	<.05
1 : 35	3 : 14	21	<.0001
2 : 23	3 : 14	9	<.20

H.4.1.3 TE at Session 20

Rank averages and totals

Level	Rank average	Rank total
1: Alphabetical	2.88	35
2: Predictive	2.00	24
3: Hybrid	1.13	14

Friedman's S = 220.5
CHI(2)= 18.38, p=0.0001 (No correction for ties.)
CHI(2)= 18.77, p=0.0001 (With correction for ties.)

==

Table of rank frequencies

	Conditions		
Rank	1	2	3
1.0	0	1	10
1.5	1	0	1
2.0	0	10	1
2.5	0	0	0
3.0	11	1	0
Rank Tot	35	24	14
Rank Ave	2.9	2.0	1.1

==

Multiple Comparisons Based on Friedman Rank-Sums

Condition A	Condition B	Absolute Difference	Probability Level
1 : 34.5	2 : 24	10.5	<.10
1 : 34.5	3: 13.5	21	<.0001
2 : 24	3: 13.5	10.5	<.10

H.4.1.4 Alphabetical TE at Session 1, 10, and 20

Rank averages and totals

Level	Rank average	Rank total
1: Alphabetical	1.92	23
2: Predictive	2.13	26
3: Hybrid	1.96	24

Friedman's S = 3.5
CHI(2)= 0.29, p=0.8643 (No correction for ties.)
CHI(2)= 0.30, p=0.8616 (With correction for ties.)

==

Table of rank frequencies

	Conditions		
Rank	1	2	3
1.0	6	2	3
1.5	0	1	1
2.0	1	5	5
2.5	0	0	0
3.0	5	4	3
RankTot	23	26	24
RankAve	1.9	2.1	2.0

Multiple Comparisons Based on Friedman Rank-Sums

Condition A	Condition B	Absolute Difference	Probability Level
1 : 23	2 : 25.5	2.5	>.20
1 : 23	3 : 23.5	.5	>.20
2 : 25.5	3 : 23.5	2	>.20

H.4.1.5 Predictive TE at Session 1, 10, and 20

Rank averages and totals

Level	Rank average	Rank total
1: Alphabetical	1.58	19
2: Predictive	2.33	28
3: Hybrid	2.08	25

Friedman's S = 42
CHI(2)= 3.50, p=0.1738 (No correction for ties.)
No ties in this data set.
==
Table of rank frequencies

	Conditions		
Rank	1	2	3
1.0	8	2	2
1.5	0	0	0
2.0	1	4	7
2.5	0	0	0
3.0	3	6	3
Rank Tot	19	28	25
Rank Ave	1.6	2.3	2.1

==
Multiple Comparisons Based on Friedman Rank-Sums

Condition A	Condition B	Absolute Difference	Probability Level
1 : 19	2 : 28	9	<.20
1 : 19	3 : 25	6	>.20
2 : 28	3 : 25	3	>.20

H.4.1.6 Hybrid TE at Session 1, 10, and 20

Rank averages and totals

Level	Rank average	Rank total
1: Alphabetical	1.67	20
2: Predictive	2.21	27
3:Hybrid	2.13	26

Friedman's S = 24.5
CHI(2)= 2.04, p=0.3603 (No correction for ties.)
CHI(2)= 2.09, p=0.3526 (With correction for ties.)

Table of rank frequencies

Rank	Conditions 1	2	3
1.0	6	3	2
1.5	0	1	1
2.0	4	2	5
2.5	0	0	0
3.0	2	6	4
Rank Tot	20	27	26
Rank Ave	1.7	2.2	2.1

==

Multiple Comparisons Based on Friedman Rank-Sums

Condition A	Condition B	Absolute Difference	Probability Level
1 : 20	2 : 26.5	6.5	>.20
1 : 20	3 : 25.5	5.5	>.20
2 : 26.5	3 : 25.5	1	>.20

H.5 Keyboard Rank

H.5.1 Friedman Test of Keyboard Rank across Sessions

Level	Rank average	Rank total
1: Alphabetical	2.93	59
2: Predictive	2.08	42
3: Hybrid	1.00	20

Friedman's S = 744.5

CHI(2)= 37.22, p=0.0000 (No correction for ties.)
CHI(2)= 37.70, p=-.0000 (With correction for ties.)

Multiple comparisons based on Friedman Rank-Sums for keyboard rank across sessions

Condition A	Condition B	Absolute Difference	Probability Level
1 : 58.5	2 : 41.5	17	<.025
1 : 58.5	3 : 20	38.5	<.0001
2 : 41.5	3 : 20	21.5	<.005

H.5.2 Friedman Test: Alphabetical Rank versus Session Blocked by Participant

S = 17.79 DF = 19 P = 0.537
S = 45.76 DF = 19 P = 0.001 (adjusted for ties)

Session	N	Est Median	Sum of Ranks
1	12	2.2000	62.0
2	12	2.9500	112.5
3	12	2.9500	104.5
4	12	2.9500	104.5
5	12	3.0000	124.5
6	12	2.9500	114.0
7	12	3.0000	144.0
8	12	3.0000	126.5
9	12	3.0000	130.5
10	12	3.0000	134.5
11	12	3.0000	134.5
12	12	3.0000	125.0
13	12	3.0000	152.5
14	12	3.0000	134.5
15	12	3.0000	134.5
16	12	3.0000	134.5
17	12	3.0000	125.0
18	12	3.0000	134.0
19	12	3.0000	144.0
20	12	3.0000	144.0

Grand median = 2.9500

H.5.3 Friedman Test: Predictive Rank versus Session blocked by Participant

S = 17.29 DF = 19 P = 0.570
S = 30.90 DF = 19 P = 0.041 (adjusted for ties)

Session	N	Est Median	Sum of Ranks
1	12	2.7250	169.5
2	12	2.3250	152.0
3	12	2.0500	134.5
4	12	2.2000	154.0
5	12	2.0250	141.0
6	12	2.1250	146.0
7	12	1.9750	110.0
8	12	2.0250	134.5
9	12	2.0000	112.0
10	12	1.9750	118.5
11	12	1.9250	110.0
12	12	2.0000	129.5
13	12	1.9250	101.5
14	12	2.0000	128.0
15	12	2.0000	120.5
16	12	1.9500	110.5
17	12	1.9250	112.5
18	12	2.0000	131.0
19	12	1.8750	93.5
20	12	1.9750	111.0

Grand median = 2.0500

H.5.4 Friedman Test: Hybrid Rank versus Session blocked by Participant

S = 7.71 DF = 19 P = 0.989
S = 19.48 DF = 19 P = 0.426 (adjusted for ties)

Session	N	Est Median	Sum of Ranks
1	12	1.2250	162.0
2	12	1.0250	116.0
3	12	1.0250	142.0
4	12	1.0250	126.0
5	12	1.0000	116.0
6	12	1.0000	116.0
7	12	1.0250	122.0
8	12	1.0000	112.0
9	12	1.0250	139.5
10	12	1.0000	119.5
11	12	1.0000	129.5
12	12	1.0000	119.5
13	12	1.0250	122.0
14	12	1.0000	109.5
15	12	1.0250	119.5
16	12	1.0250	129.5
17	12	1.0250	139.5
18	12	1.0000	116.0
19	12	1.0250	142.0
20	12	1.0250	122.0

Grand median = 1.0250

H.5.5 Participants' Comments Associated with Ranking Layouts

H.5.5.1 Session 1

Participant	Alphabetical	Predictive	Hybrid
1	Familiar, but so much scrolling to get anywhere	There is help to minimize travel.	Tough to get used to it
2	Knew where everything was	A little harder. Need to do too much searching	The predictive set did not always have the letter I needed
3	I liked it because it was not predictive. I know where the letters are going to be. Would prefer if the space key was in the middle.	I had to think too much with this one. Had to keep searching for the letters.	It wasn't too bad. I would prefer less predictive keys (~5) to allow for quicker scanning of that set.
4	Tedious	Sometimes pass letters	Hate Caps key location and that it requires two keystrokes to get to the Delete key
5	Easy. I had an idea of where I needed to go	I could not find anything. I missed letters. I just could not find the letters.	Switching between the predictive keys and the alphabetical row was complicated… preferred just going to the top row.
6	--	--	--
7	Had to travel too far to get to a letter	Not consistent.	Did not like Caps key location, but starting in the middle makes sense
8	I know where the letters and can start immediately moving towards letter	I can't forward think…	Distracting that the letters jumped around
9	--	--	--
10	Familiar	Keys constantly on the move, needed to search for letters	Easy
11	Quickly found the letter needed and direction to travel	Mentally taxing… dynamic layout requires too much focus	Hybrid was easiest. Space in the middle is good
12	No learning curve. No concentration required.	If performance is much better when all combinations are learned, then it may be worth it.	Requires higher concentration and focus.

H.5.5.2 Session 10

Participant	Alphabetical	Predictive	Hybrid
1	Big problem of overshooting using typematic. Easy to make mistakes	Miss the letters	Seems to be faster
2	Most effort.	Fastest with this one	Some inefficiency at starting. Don't know where to move.
3	--	--	--
4	--	--	Placement of Cap key is bothersome
5	Takes the longest	Sometimes I miss the letters	If the letter I need is not in the predictive keys, I at least know where else to get it
6	Too much travel	More demanding	Hybrid is more predictive than the Predictive keyboard
7	More finger movement	Very demanding	Mental aspect of Predictive and physical work of alphabetical… I never know where to go
8	Keep moving all the time	Letters are always scrambled	Easy to learn letter locations, shorter distance to travel
9	Requires least focus, but too many key presses	Finding letters is a lot of work	Less focus on searching, but most difficult to use when correcting errors
10	Repetitive motion	Always trying to figure out where the letters will be	Predictive set is smaller, easier to use
11	Not too much thinking involved	Requires too much visual focus and certain letters never appear where you would expect them	Easier scanning
12	This is the best because there is no scanning	Intensive scanning	Same as Predictive, but not as intense

H.5.5.3 Session 20

Participant	Alphabetical	Predictive	Hybrid
1	Generally know where to travel, but too much finger movement required	Target usually appears in first 7-8 keys, but I tend to overshoot	Hybrid is fastest and most accurate
2	Boring. I hate this keyboard. Takes too long to get to the letter.	Constantly adjusting my eyes to search for the letters.	Has predictive and regular set. I prefer looking at multiple rows… like in a standard keyboard.
3	Frustrating. Spend too much time waiting for the cursor to get to the letter.	Spend too much time looking for the letter and would often pass (overshoot) the letter.	Had to use more fingers with this one, but still easier.
4	Can't reach greater speeds	Must actively be thinking about optimal travel path	My favorite
5	Takes too long	Too overwhelming	The predictive keys usually had the letter I needed
6	Always trying to determine how to travel faster	More scanning involved	Type more quickly
7	It is always clear as to where the letters are	Trying to anticipate changing keyboard layouts and too much finger movement	Better ability to recognize changing layouts. Less movement too
8	Too much finger work	Too much scanning, strains my eyes	Good combination
9	More key presses and made more errors	Too much scanning	Type fastest with Hybrid
10	Dull	Not as predictive as I would hope	My favorite. I know where the letters are if they are not in the predictive set
11	Boring! Very time consuming	A lot of physical effort and need precision in movement	More efficient layout
12	Typematic keying should be faster for this keyboard only.	Scanning is a lot harder. More visual search and more mental work. Not confident on this one.	Does not combine benefits of others. Scanning is hard and need to double check to be sure that you did not miss the letter in the predictive set.

H.6 Keyboard Rating

H.6.1 Linear Mixed Effects Model Results

Data: Ratings.Data

AIC	BIC	logLik
-0.08051581	86.73963	19.04026

Random effects:
Formula: ~ Keyboard + Session | Subject
Structure: General positive-definite

	StdDev	Corr		
(Intercept)	0.38251716	(Intr)	KyFOCL	KybHYB
KeyboardFOCL	0.38997575	-0.871		
KeyboardHYB	0.61194149	-0.888	0.948	
Session	0.01698603	-0.200	0.138	0.116
Residual	0.20591681			

Correlation Structure: Continuous AR(1)
Formula: ~ 1 | Subject
Parameter estimate(s):
Phi
0.06562147

Fixed effects: log(Rating) ~ Keyboard + Session + Session^2 + Session * Keyboard

	Value	Std.Error	DF	t-value	p-value
(Intercept)	1.092360	0.1157348	702	9.438474	<.0001
KeyboardFOCL	0.302923	0.1187626	702	2.550656	0.0110
KeyboardHYB	0.053054	0.1806252	702	0.293726	0.7691
Session	-0.029663	0.0079068	702	-3.751533	0.0002
I(Session^2)	0.000906	0.0002750	702	3.294509	0.0010
KeyboardFOCLSession	-0.005384	0.0031604	702	-1.703612	0.0889
KeyboardHYBSession	-0.009172	0.0031453	702	-2.916095	0.0037

Standardized Within-Group Residuals:

Min	Q1	Med	Q3	Max
-2.642053	-0.5364349	-0.06468881	0.469129	5.127217

Number of Observations: 720
Number of Groups: 12

Analysis of Variance Table

	numDF	denDF	F-value	p-value
(Intercept)	1	702	451.8950	<.0001
Keyboard	2	702	26.2031	<.0001
Session	1	702	9.1923	0.0025
I(Session^2)	1	702	10.7396	0.0011
Session:Keyboard	2	702	4.2899	0.0141

Approximate 95% confidence intervals

Fixed effects:

	lower	est.	upper
(Intercept)	0.8651318816	1.0923596401	1.319587399
KeyboardFOCL	0.0697500667	0.3029225215	0.536094976
KeyboardHYB	-0.3015760115	0.0530543405	0.407684693
Session	-0.0451867089	-0.0296628056	-0.014138902

I(Session^2)	0.0003660619	0.0009059723	0.001445883
KeyboardFOCLSession	-0.0115891806	-0.0053841498	0.000820881
KeyboardHYBSession	-0.0153475585	-0.0091721388	-0.002996719

Random Effects:
Level: Subject

	lower	est.	upper
sd((Intercept))	0.24793915	0.38251716	0.59014229
sd(KeyboardFOCL)	0.25396280	0.38997575	0.59883212
sd(KeyboardHYB)	0.40084708	0.61194149	0.93420263
sd(Session)	0.01076694	0.01698603	0.02679731
cor((Intercept),KeyboardFOCL)	-0.96237862	-0.87103195	-0.60278531
cor((Intercept),KeyboardHYB)	-0.96717561	-0.88800845	-0.65170873
cor((Intercept),Session)	-0.76401950	-0.19951591	0.53800514
cor(KeyboardFOCL,KeyboardHYB)	0.82093474	0.94824593	0.98575019
cor(KeyboardFOCL,Session)	-0.58161557	0.13847273	0.73688877
cor(KeyboardHYB,Session)	-0.59217794	0.11595243	0.72303048

Correlation structure:

	lower	est.	upper
Phi	0.01939011	0.06562147	0.19964

Within-group standard error:

lower	est.	upper
0.1951001	0.2059168	0.2173332

anova(RatingData, L=c(KeyboardFOCL=1, KeyboardHYB=-1))
F-test for linear combination(s)

KeyboardFOCL	KeyboardHYB
1	-1

	numDF	denDF	F-value	p-value
1	1	702	8.2324	0.0042

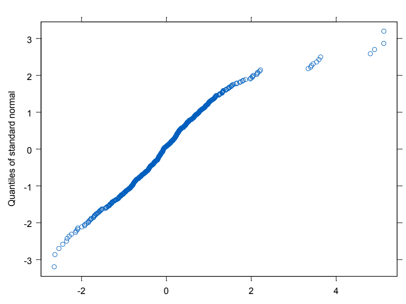
anova(RatingData, L=c(KeyboardFOCLSession=1, KeyboardHYBSession=-1))
F-test for linear combination(s)

KeyboardFOCLSession	KeyboardHYBSession
1	-1

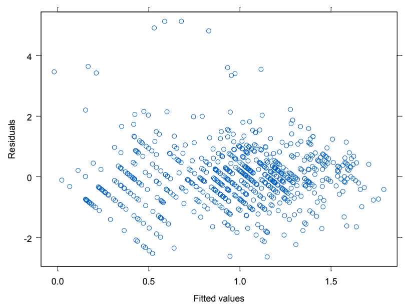
	numDF	denDF	F-value	p-value
1	1	702	1.450372	0.2289

H.6.2 Ratings Test of Model Assumptions

H.6.2.1 Test of Normality



H.6.2.2 Test of Homogeneity of Variance



H.6.3 Rating Pairwise t-tests

Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
					95% Confidence Interval of the Difference				
		Mean	Std. Deviation	Std. Error Mean	Lower	Upper			
Pair 1	Alphabetical - Predictive	-.68067	.99089	.28605	-1.31025	-.05108	-2.380	11	.037
Pair 2	Predictive - Hybrid	.79167	1.18306	.34152	.03999	1.54334	2.318	11	.041
Pair 3	Hybrid - Alphabetical	-.11100	1.70021	.49081	-1.19127	.96927	-.226	11	.825

H.7 Rank vs. Rating

Pearson correlation of Rating_ABC and ABC Rank = 0.294
P-Value = 0.000

Pearson correlation of Rating_HYB and HYB Rank = 0.473
P-Value = 0.000

Pearson correlation of Rating_FOCL and FOCL Rank = 0.544
P-Value = 0.000