HUMAN-COMPUTER INTERACTION, 2002, Volume 17, pp. 131–XXX

Empirical Bi-Action Tables: A Tool for the Evaluation and Optimization of Text-Input Systems. Application I: Stylus Keyboards

Dominic Hughes, James Warren, and Orkut Buyukkokten
Stanford University

ABSTRACT

We introduce a technique that, given any text input system *A* and novice user *u*, will predict the peak expert input speed of *u* on *A*, avoiding the costly process of actually training *u* to expert level. Here, *peak* refers to periods of ideal performance, free from hesitation, or concentration lapse and *expert* refers to asymptotic competence (e.g., touch typing, in the case of a two-handed keyboard). The technique is intended as a feedback mechanism in the interface development cycle between abstract mathematical modeling at the start (Fitts' law, Hick's law, etc.) and full empirical testing at the end.

Dominic Hughes is a computer scientist specializing in logic and mathematical foundations of computation; he is a Research Associate in the Computer Science Department of Stanford University. **James Warren** is a computer scientist with interests in machine learning and optimization; he is a PhD student in the Scientific Computing and Computational Mathematics Program of Stanford University. **Orkut Buyukkokten** is a computer scientist with interests in databases and human–computer interaction; he is a PhD student in the Computer Science Department of Stanford University.

CONTENTS

The utility of the technique in iterative design is contingent on what we call the monotonicity principle: For each user *u*, if our prediction of peak expert input speed for *u* is higher on system *A* than on system *B*, continuous text input by *u* after training to expert level will be faster on *A* than on *B*. Here, *continuous* refers to actual real-world use, subject to errors, physical fatigue, lapses of concentration, and so forth. We discuss the circumstances under which monotonicity is valid.

The technique is parametric in the character map—that is, in the map from actions (keystrokes, gestures, chords, etc.) to characters. Therefore, standard heuristic algorithms can be employed to search for optimal character maps (e.g., keyboard layouts). We illustrate the use of our technique for evaluation and optimization in the context of stylus keyboards, first benchmarking a number of stylus keyboards relative to a simple alphabetic layout and then implementing an ant algorithm to obtain a machine-optimized layout.

1. INTRODUCTION

As with any design process, the design of text input systems requires a feedback mechanism to iterate to better solutions. Full empirical evaluation of expert performance on a new text input system is costly, due to the vast number of hours required to train test participants to expert level. Furthermore, evaluation is highly sensitive: Even a minor modification in the system (e.g., switching the position of a key on a keyboard layout, changing a chord on a chording keyboard, or modifying a gesture of a glove input language) forces a repeat of all experiments because participants must be retrained to expert level on the modified system. Such high cost and sensitivity render the iterative design of text input systems impractical; interface design becomes more of an art than a science.

In certain cases, rules, models, or equations (e.g., Fitts' law, Hick's law, and the power law; Card, Moran, & Newell, 1983) can be used to generate a feedback loop in the early stages of the design process. However, these techniques are not without drawbacks. First, by the very nature of abstraction, there can be problems of fidelity and resolution (Noel & McDonald, 1989). Second, there is the problem of lack of generality: Some systems may be beyond the scope of laws. For example, how does one model intricate gestures with a glove or stylus, or the complex parallelism and interference between fingers at a two-handed keyboard?

We introduce a technique intended as a tool in the interface design cycle between abstract mathematical modeling at the start (Fitts' law, etc.) and full empirical user testing at the end. Given any text input system *A* and novice user *u*, the technique predicts the peak expert input speed of *u* on *A*, avoiding the costly process of actually training *u* up to expert level. Here, *peak* refers to periods of ideal performance, free from hesitation or concentration lapse, and *expert* refers to asymptotic competence (e.g., touch typing in the context of a two-handed keyboard).

Our conceptual starting point is a strict separation of the text input system into two parts:

- *Physical aspect*: The *actions* performable on the device—for example, a keystroke of a two-handed keyboard, the articulation of a gesture with a stylus or glove, or the depression of a chord on a two-handed keyboard.
- *Logical aspect*: The *character map*, specifying the interpretation of each action as a character—for example, "striking the top left key" maps to *Q*, "a vertical down-stroke with the index finger of the glove" maps to *I*, "chording the two outermost keys" maps to *Y*.

For each (potentially novice) user *u* and text input system *A* we capture the pure physical aspect of interaction between *u* and *A*, in total isolation from the logical aspect, as an empirical bi-action table *E*. For each pair of actions *i* and *j* (e.g., keystrokes on a two-handed keyboard, gestures with a stylus or glove, etc.), the table entry E_{ij} is the aggregate result of experiments recording the time to complete j having just completed i, in the absence of any character map. Then, given an arbitrary character map, we obtain a prediction of peak expert input speed by using a table of character bi-gram frequencies (see Figure 1).

It is crucial that the experiments are completely independent of the logical aspect (i.e., are conducted in the absence of a character map). For example, in the context of glove gestures we might ask a user to form a fist then flatten the hand—without reference to any particular interpretation of these actions as characters. In the case of stylus keyboards, we might place users at a completely blank grid of keys and ask them to tap the top left key followed by the bottom right key. The reasoning behind our choice to conduct pure physical experimental trials in the absence of a character map is as follows:

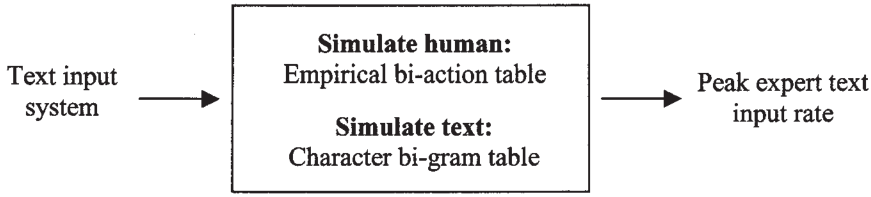
1. The experimental data capture physical coordination representative of expert-level production of text during peak concentration. Pairs of actions are executed fluently, as they would be by an expert user who is completely familiar with an ambient character map—without actually having to train the user to expert level on a character map.
2. Once the empirical bi-action table is obtained, we can immediately predict peak expert input speed for the device under any character map.
3. Because predictions are obtained immediately for any character map, we can employ standard heuristic algorithms to search for an optimal character map (e.g., keyboard layout).

Our prediction of peak expert input speed is not intended to be an estimate of real continuous performance. The latter is subject to errors, physical fatigue, and lapses in concentration, factors that are highly unpredictable, varying not only between users but also on a session-by-session basis for a single user. However, in the absence of realistic predictions of continuous input speeds, our technique is nonetheless a useful feedback mechanism for the iterative design of text-input systems if one accepts the validity of the *monotonicity*[1] *principle,* which we present in a strong and a weak form:

> *Strong Monotonicity Principle*: For any user u and any text input systems A and B, if our prediction of peak expert input speed for u is higher on A than on B, continuous text input by u after training to expert level will be faster on A than on B.

1. The terminology derives from mathematics. A function $f: X \to Y$ between ordered sets X and Y is monotonic if and only if $\forall\ x, x' \in X.\ [x \leq x' \Rightarrow fx \leq fx']$.

Figure 1. **Predicting peak expert text input rates.**



> *Weak Monotonicity Principle*: For any user *u* and text input systems *A* and *B* that differ only in character map (and hence have the same empirical bi-action table), if our prediction of peak expert input speed for *u* is higher on system *A* than on system *B*, continuous text input by *u* after training to expert level will be faster on *A* than on *B*.

In Section 2 we outline why the strong form does not hold in general, particularly when *A* and *B* are physically very dissimilar. Therefore, when using our predictions of peak expert input speed to benchmark one text input system *A* against a very different system *B* (e.g., Morse code against handwriting recognition), careful consideration is required on the part of the researcher before the results can be deemed meaningful.

Also in Section 2, we argue that the weak monotonicity principle holds. Consequently, it is valid to use the empirical bi-action technique to evaluate and compare alternative character maps for the same device and to search for optimal character maps using heuristic algorithms. We demonstrate this approach in Section 3 in the context of stylus keyboard layouts. We benchmark a number of stylus keyboards relative to a simple alphabetic layout, then implement an ant algorithm to obtain a machine-optimized layout. In particular, we validate our technique by successfully correlating predicted peak expert input rates with previous results on two stylus keyboards: the OPTI (MacKenzie & Zhang, 1999) and the FITALY (Textware™ Solutions, 2000). We summarize this illustrative example later.

In Section 4, we discuss the strengths and weaknesses of our technique relative to full empirical testing and abstract mathematical modeling (Fitts' law, etc.). Having illustrated the approach in the context of stylus keyboards, we note that empirical bi-action tables could be used in the analysis and design of a wide variety of text input systems, such as two-handed keyboards, chording keyboards, cell phones, glove gesture input, and forms of stylus input including Graffiti®, Quikwriting (Perlin, 1998), and Unistrokes (Goldberg & Richardson, 1993).

In Section 3, we illustrate our technique in the context of stylus keyboard design. A stylus keyboard is a graphical keyboard displayed on a touch screen, on which users type by tapping with a stylus (pen). An example is the pop-up

QWERTY available on many personal digital assistants (PDAs). Other methods of text entry with a stylus include handwriting recognition and gesture recognition.

It is a point of contention as to whether users of mobile devices spend enough time entering text to be willing to expend the effort to acquire expertise with a faster layout than the familiar QWERTY. We justify the search for faster layouts as follows: (a) There is a market, as witnessed by sales of a commercial layout called the FITALY (Textware Solutions, 2000); (b) although a large volume of text will rarely be entered in a single session, high numbers of short messages are likely; and (c) only *after* researchers have explored the space of optimized keyboards and have understood what can be gained by switching from QWERTY can we conclude that a majority of users would continue using the QWERTY. We do not dwell on these issues in this article because we are using stylus keyboard design as a domain to illustrate a more general approach.

Both full empirical testing and abstract mathematical modeling have been used in stylus keyboard design (Hunter, Zhai, & Smith, 2000; Lewis, LaLomia, & Kennedy, 1999; MacKenzie, Nonnecke, McQueen, Riddersma, & Meltz, 1994; MacKenzie & Zhang, 1999; MacKenzie, Zhang, & Soukoreff, 1999; Soukoreff & MacKenzie, 1995). Abstract approaches use an equational characterization of human motion (known as Fitts' law; Fitts, 1954) to simulate user input at keyboards and hence obtain estimates of text entry rate. Then, with an evaluation function at hand, one can apply off-the-shelf techniques to find optimized keyboard layouts. For example, Hunter et al. (2000) employed dynamic simulation and the Metropolis method (see also Zhai, Hunter, & Smith, 2000). A drawback of the pure analytical approach is that there is experimental evidence (Section 2.2 of MacKenzie, 1991) that small-scale hand motions are not accurately characterized by the law.

Due to the inherently abstract nature of Fitts' law, previous works have only considered the distance between two keys as a predictor of the duration of the motion between these keys. We show that this duration depends also on the first key position and on the relative position of the second key. These dependencies, as well as any other more subtle dependencies that could be impossible to model, are automatically taken into account by our empirical bi-action table, supporting its use in the intermediate design phase between an initial use of laws and the final full user testing.

In the context of stylus keyboards, an action is a tap of the stylus, so in Section 3 we refer to bi-taps instead of bi-actions. In Section 3.9 we describe the implementation of an ant algorithm to find an optimized stylus keyboard layout, and in Section 3.11 we benchmark a number of keyboards against a naive alphabetical layout ABC. The peak input rate of the layout produced by the ant algorithm was 15.65% faster than the ABC, the FITALY was 13.35% faster, the OPTI was 11.63% faster, and a variant of the ABC with a center Space key was 3.95% faster. Until a larger corpus of bi-tap data has been amassed, these figures should not be considered final because only five participants were tested in the original construction of the bi-tap table. Furthermore, the output of the ant algorithm had the advantage of the coincidence of training data and test data.

2. EMPIRICAL BI-ACTION TABLES

Given a text input system S, we perform experiments to capture the physical aspect of S in an empirical bi-action table E: For each pair of actions i and j (e.g., keystrokes on a two-handed keyboard, gestures with a stylus or glove, etc.), the entry E_{ij} is the aggregate result of experiments recording the time to complete j having just completed i. For the reasons outlined in Section 1, it is crucial that the experiments be conducted in complete isolation from the logical aspect (i.e., completely independently of any character map). An illustration of how to perform such experiments is given in detail in Section 3, in the context of stylus keyboards.

Given a character map K (i.e., an assignment of actions to characters), a table of bi-gram probabilities P and an empirical bi-action table E,

Character map	K: character a $\rightarrow$ action $K(\alpha)$
Bi-grams	P: (character α, character β) $\rightarrow$ probability $P(\alpha, \beta)$
Bi-actions	E: (action i, action j) $\rightarrow$ duration $E(i, j) = E_{ij}$

the peak expert text input rate $R(K, P, E)$, in characters per unit time, is given by:

$$R(K,P,E) = \frac{1}{\sum_{\alpha,\beta} P(\alpha,\beta) E(K(\alpha), K(\beta))} \tag{1}$$

where α and β range over the character set.

This equation can be decomposed and understood as follows:

(α, β) *bi-gram execution time*	$M_{P,E}(\alpha,\beta) = E(K(\alpha), K(\beta))$
Mean bi-gram execution time	$\bar{M}(K,P,E) = \sum_{\alpha,\beta} P(\alpha,\beta) M_{P,E}(\alpha,\beta)$
Peak expert text-input rate	$R(K,P,E) = \frac{1}{\bar{M}(K,P,E)}$

$\bar{M}$ *(K, P, E)* is the mean time taken to input an ordered pair of characters (bi-gram) under the character map K, with text represented by the bi-gram probabilities P and motion modeled by the empirical bi-action table E. $M_{P,E}(\alpha, \beta)$ is the time taken to input the ordered pair of characters (α, β) under the character map K, with motion modeled by the empirical bi-action table E. Note that this depends only on the character map K and empirical bi-action table E.

From this decomposition we observe that our Equation 1 for text-input rate $R(K, P, E)$ is similar to Equation 5 of MacKenzie et al. (1999) with (α, β) bi-gram movement time set to $E(K(\alpha), K(\beta))$.

The peak expert text input rate defined earlier is not intended to be an estimate of real continuous performance. Actual text input, being subject to factors such as physical fatigue and lapses in concentration, will only reach peak rate in short bursts. The peak expert text input rate is thus a useful mechanism for the iterative design of text input systems only if one accepts the validity of the *monotonicity principle*, which we present in a strong and a weak form:

> *Strong Monotonicity Principle*: For any user u and any text input systems A and B, if our prediction of peak expert text input rate for u is higher on A than on B, continuous text input by u after training to expert level will be faster on A than on B.
>
> *Weak Monotonicity Principle*: For any user u and text input systems A and B that differ only in character map (and hence have the same empirical bi-action table), if our prediction of peak expert text input rate for u is higher on system A than on system B, continuous text input by u after training to expert level will be faster on A than on B.

The strong form does not hold in general, particularly when A and B are physically very dissimilar. The sources of fatigue may be very different for different devices, as may be concentration levels required during use. Consider, for example, the case of A, an optimized gesture language such as Unistrokes (Goldberg & Richardson, 1993), and B, a hunt-and-tap stylus keyboard. Due to the continuous visual scanning required during input at the stylus keyboard, actual sustained input on B may be more prone to fatigue or errors, so a slightly higher prediction of peak expert input speed for B than for A may not accurately represent better real-world sustained performance on B than on A. Therefore when using our predictions of peak expert input speed to benchmark one text input system A against a very different system B, careful consideration is required on the part of the researcher before the results can be deemed meaningful.

However, when A and B use the same device and set of actions and differ only in character map, factors such as concentration laps and fatigue will be similar. Therefore, continuous text entry rate will be a similar dilution of peak text entry rate in each case, and the Weak Monotonicity Principle holds. Consequently, it is valid to use the empirical bi-action technique to evaluate and compare alternative character maps for the same device and to search for optimal character maps using heuristic algorithms. We demonstrate this approach in Section 3 in the context of stylus keyboard layouts.

3. ILLUSTRATIVE EXAMPLE: STYLUS KEYBOARDS

We illustrate the use of an empirical bi-action table in the context of stylus keyboard evaluation and optimization. We describe in detail the experiment to generate the empirical bi-action table, benchmark various layouts against a simple alphabetical layout, and then we implement an ant algorithm to search for an optimal layout. We validate our technique by correlating predicted peak expert input speeds with previous results on the OPTI and FITALY layouts. Because actions are taps with a stylus, in this section we refer to bi-actions as bi-taps.

3.1. Method

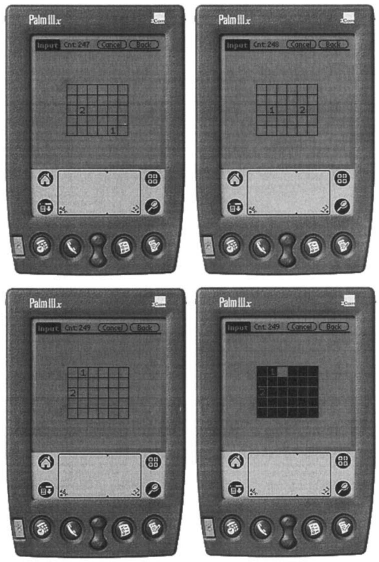
Participants. There were five participants, each right-handed. All were students, familiar with desktop computing. Four were men. The youngest was 20, the oldest was 32, and the mean age was 26.3. Of the 5, only 1 had previous experience with stylus text input. All were paid for their participation in the study.

Apparatus. For the experiments we used several different PDAs: Palm™ III, Palm IIIx, Palm VII (Palm, Inc.), and Visor (Handspring™). The software was written in C++ using CodeWarrior™ for Palm OS® (Metrowerks). The time measurements were gathered using the device clock, in terms of *time ticks*. The operating system has a GetTimeTicks() function that gives the time elapsed in milliseconds. This function is called each time the user taps on the screen, and the time elapsed between two pen taps is measured by taking the difference between two consecutive values. The data were saved in a Palm OS database and later downloaded to a desktop computer for analysis.

Procedure. Each participant underwent five separate tests. The duration of a test was approximately 30 min. Participants worked on a 5-row, 6-column blank keyboard on a PDA (Figure 2). The dimensions of the grid were as follows: width = 3.15 cm, height = 2.65 cm, square key width/height = 5.25 mm.

Define a *bi-tap* to be any ordered pair of keys (k_1, k_2). A test consisted of presenting a participant with all 900 possible bi-taps in random order. Each bi-tap was presented by labeling two of the blank keys, 1 and 2. The participant tried as

Figure 2. **Screen shots from the bi-tap experiment. The screen shots are in order of bi-taps presented to a user during the test (observe the count in the top of the display). The lower right shot shows the irritating 3-sec screen-lock incurred by committing an**



quickly as possible to tap key 1 followed by key 2. We recorded the time of each such bi-tap (i.e., the time interval between tapping key 1 and tapping key 2).

Participants were aware of the fact that their time to find and hit key 1 is not being recorded. Furthermore, they were instructed to absorb the positions of both keys before undertaking any physical motion, so that the recorded interval does not include scanning time for key 2. Should a participant be interrupted or distracted after striking key 1, they are instructed to the cancel button to be presented with the same bi-tap a second time. When an incorrect key was tapped the application emitted an ugly BEEP, and the screen locked for 3 sec. The irritation ensured an extremely low error rate.

An important aspect of the test was that all 900 bi-grams were carried out in succession rather than in isolation. This was to simulate the fact that during real typing, participants typically adopt a natural rest position with their wrist on the side of the PDA. For consistency, all participants were asked to work the PDA in hand rather than supported flat on a desk.

3.2. Results

We adopt the chess naming convention for keys as depicted in Figure 3, row index A to F from left to right, column index 1 to 5 from bottom to top.

Figure 3. **Naming convention for keys.**

A5	B5	C5	D5	E5	F5
A4	B4	C4	D4	E4	F4
A3	B3	C3	D3	E3	F3
A2	B2	C2	D2	E2	F2
A1	B1	C1	D1	E1	F1

The results of the experiment, aggregated across the 25 tests of the 5 participants, are shown in Figure 4, which we refer to as the empirical bi-tap table. The entry at row k_1 and column k_2 represents the time taken by a generic user between tapping key k_1 and key k_2. The largest entry in the table is .330 sec, for the long diagonal bi-tap A1 to F5. The smallest entry is .147, for the double-tap of F3.

Note that we use $\bar{E}$ to denote the empirical bi-tap table of Figure 4. As detailed in the next section, the overline is to remind us that $\bar{E}$ is the mean of five bi-tap tables, one per test participant. We write $\bar{E}(k_1, k_2)$ for the entry at row k_1 and column k_2.

3.3. Calculating the Entries of the Empirical Bi-Tap Table

Having undergone the experiment five times, each user generated five data points per bi-tap; hence, we have a total of 25 data points per bi-tap. During the testing procedure we occasionally observed lapses of concentration by a user in the middle of a bi-tap. Such instances result in anomalous data points that are not in agreement with our objective of capturing the purely physical "minimum transition time" between a pair of taps.

Outliers were discarded uniformly with the following procedure: For each user, and for each bi-tap, discard the data points that are more than twice the duration of the minimum bi-tap time in the user's quintuple of recorded data points. From a total set of 21,750 data points for bi-taps consisting of distinct keys, this procedure discards 877 points (i.e., an outlier cutoff of 4%). Note that we remove outliers on a per user basis to allow for the fact that some users are uniformly faster and more coordinated than others.

Figure 4. **Empirical bi-tap table . The value (k_1, k_2) at row k_1 and column k_2 is the time taken to tap key k_1 followed by key k_2 in seconds. (See Figure 3 for the naming convention for the keys on the 5 × 6 grid.) The value of (k_1, k_2) is the mean of the corresponding values in the bi-tap tables of the 5 participants. Section 3.3 explains how the value is computed from the 25 experimental data points for (k_1, k_2).**

1st/2nd	A1	B1	C1	D1	E1	F1	A2	B2	C2	D2	E2	F2	A3	B3	C3
A1	0.157	0.157	0.188	0.225	0.254	0.281	0.165	0.175	0.203	0.234	0.245	0.296	0.197	0.2	0.221
B1	0.176	0.153	0.165	0.181	0.227	0.251	0.178	0.169	0.173	0.204	0.223	0.256	0.204	0.188	0.202
C1	0.202	0.172	0.154	0.158	0.194	0.214	0.203	0.172	0.17	0.175	0.216	0.227	0.237	0.202	0.194
D1	0.232	0.204	0.173	0.151	0.161	0.187	0.241	0.21	0.175	0.165	0.17	0.203	0.241	0.233	0.2
E1	0.265	0.227	0.2	0.164	0.163	0.16	0.269	0.237	0.21	0.173	0.162	0.166	0.267	0.241	0.222
F1	0.285	0.268	0.238	0.199	0.166	0.156	0.286	0.27	0.237	0.204	0.17	0.163	0.278	0.269	0.242
A2	0.167	0.172	0.195	0.219	0.242	0.273	0.154	0.157	0.189	0.227	0.257	0.263	0.164	0.171	0.196
B2	0.17	0.164	0.16	0.181	0.231	0.223	0.17	0.156	0.159	0.202	0.224	0.259	0.174	0.164	0.17
C2	0.202	0.179	0.16	0.164	0.182	0.216	0.206	0.173	0.149	0.156	0.191	0.216	0.201	0.18	0.165
D2	0.23	0.208	0.17	0.159	0.168	0.184	0.232	0.203	0.167	0.156	0.156	0.188	0.234	0.204	0.168
E2	0.263	0.234	0.202	0.167	0.164	0.167	0.272	0.227	0.202	0.167	0.158	0.167	0.268	0.232	0.204
F2	0.286	0.26	0.236	0.2	0.174	0.164	0.268	0.259	0.221	0.2	0.165	0.156	0.289	0.27	0.228
A3	0.189	0.193	0.2	0.219	0.246	0.274	0.17	0.172	0.188	0.228	0.242	0.267	0.153	0.168	0.186
B3	0.199	0.194	0.186	0.194	0.232	0.26	0.17	0.167	0.163	0.199	0.214	0.24	0.164	0.155	0.166
C3	0.213	0.199	0.197	0.198	0.197	0.228	0.2	0.171	0.164	0.162	0.186	0.214	0.197	0.168	0.156
D3	0.227	0.205	0.196	0.19	0.184	0.196	0.234	0.204	0.174	0.157	0.164	0.188	0.23	0.191	0.163
E3	0.271	0.236	0.204	0.191	0.187	0.194	0.26	0.238	0.203	0.173	0.163	0.173	0.26	0.225	0.196
F3	0.293	0.259	0.257	0.225	0.206	0.198	0.284	0.257	0.242	0.204	0.177	0.16	0.275	0.249	0.231
A4	0.227	0.224	0.228	0.226	0.261	0.272	0.192	0.186	0.199	0.227	0.245	0.307	0.161	0.172	0.187
B4	0.227	0.215	0.218	0.234	0.25	0.259	0.206	0.186	0.187	0.197	0.226	0.244	0.176	0.169	0.166
C4	0.242	0.216	0.216	0.216	0.22	0.242	0.215	0.2	0.184	0.187	0.204	0.237	0.2	0.176	0.157
D4	0.26	0.247	0.214	0.205	0.21	0.223	0.25	0.219	0.199	0.187	0.19	0.205	0.23	0.206	0.17
E4	0.285	0.247	0.24	0.212	0.222	0.21	0.264	0.223	0.225	0.204	0.188	0.198	0.259	0.23	0.199
F4	0.308	0.285	0.25	0.236	0.227	0.223	0.292	0.271	0.244	0.222	0.195	0.188	0.276	0.258	0.241
A5	0.24	0.244	0.257	0.251	0.264	0.284	0.219	0.23	0.228	0.241	0.273	0.279	0.19	0.186	0.204
B5	0.241	0.234	0.227	0.258	0.25	0.278	0.225	0.221	0.222	0.219	0.231	0.268	0.191	0.186	0.196
C5	0.242	0.25	0.242	0.251	0.237	0.258	0.242	0.224	0.221	0.219	0.221	0.246	0.219	0.194	0.182
D5	0.279	0.262	0.246	0.229	0.244	0.242	0.264	0.235	0.221	0.213	0.222	0.224	0.241	0.219	0.2
E5	0.286	0.277	0.259	0.257	0.233	0.225	0.287	0.265	0.224	0.231	0.208	0.213	0.267	0.236	0.23
F5	0.301	0.289	0.264	0.276	0.249	0.24	0.298	0.292	0.25	0.252	0.228	0.225	0.292	0.263	0.251

1st/2nd	D3	E3	F3	A4	B4	C4	D4	E4	F4	A5	B5	C5	D5	E5	F5
A1	0.229	0.269	0.282	0.215	0.235	0.258	0.273	0.283	0.305	0.248	0.27	0.283	0.309	0.326	0.33
B1	0.212	0.243	0.24	0.23	0.222	0.228	0.243	0.259	0.3	0.25	0.253	0.255	0.288	0.28	0.278
C1	0.206	0.216	0.241	0.253	0.226	0.229	0.225	0.237	0.249	0.27	0.259	0.271	0.256	0.255	0.279
D1	0.193	0.193	0.219	0.263	0.248	0.231	0.226	0.221	0.228	0.273	0.262	0.256	0.26	0.244	0.271
E1	0.205	0.19	0.202	0.289	0.254	0.245	0.225	0.224	0.227	0.283	0.275	0.279	0.238	0.253	0.249
F1	0.222	0.207	0.196	0.301	0.281	0.249	0.243	0.236	0.236	0.29	0.303	0.303	0.267	0.26	0.233
A2	0.244	0.255	0.276	0.196	0.2	0.221	0.252	0.281	0.261	0.232	0.232	0.241	0.258	0.283	0.31
B2	0.198	0.228	0.264	0.198	0.191	0.2	0.218	0.242	0.26	0.238	0.21	0.228	0.235	0.264	0.269
C2	0.163	0.194	0.222	0.223	0.193	0.204	0.194	0.22	0.236	0.236	0.235	0.221	0.214	0.232	0.26
D2	0.168	0.176	0.205	0.25	0.224	0.203	0.188	0.195	0.217	0.252	0.249	0.229	0.221	0.224	0.229
E2	0.17	0.171	0.174	0.275	0.244	0.226	0.197	0.196	0.205	0.286	0.264	0.242	0.225	0.215	0.236
F2	0.216	0.184	0.168	0.288	0.272	0.236	0.226	0.207	0.192	0.275	0.269	0.266	0.247	0.227	0.216
A3	0.226	0.253	0.284	0.17	0.172	0.201	0.23	0.274	0.273	0.195	0.2	0.232	0.23	0.257	0.275
B3	0.19	0.233	0.243	0.167	0.164	0.173	0.197	0.23	0.264	0.203	0.201	0.189	0.221	0.231	0.253
C3	0.164	0.187	0.212	0.211	0.173	0.164	0.168	0.205	0.226	0.226	0.198	0.195	0.198	0.213	0.235
D3	0.16	0.166	0.19	0.243	0.217	0.174	0.16	0.172	0.186	0.241	0.215	0.194	0.202	0.193	0.222
E3	0.172	0.158	0.166	0.256	0.232	0.21	0.182	0.166	0.176	0.256	0.235	0.217	0.209	0.199	0.197
F3	0.2	0.167	0.147	0.279	0.27	0.229	0.204	0.179	0.168	0.299	0.278	0.238	0.225	0.202	0.204
A4	0.213	0.245	0.272	0.152	0.164	0.195	0.222	0.254	0.277	0.166	0.178	0.21	0.222	0.266	0.277
B4	0.192	0.23	0.233	0.174	0.155	0.164	0.193	0.216	0.237	0.177	0.166	0.173	0.186	0.222	0.261
C4	0.162	0.191	0.218	0.21	0.173	0.157	0.159	0.187	0.218	0.194	0.183	0.173	0.165	0.197	0.201
D4	0.159	0.16	0.194	0.226	0.199	0.171	0.151	0.156	0.187	0.231	0.216	0.177	0.162	0.172	0.197
E4	0.174	0.16	0.17	0.251	0.224	0.204	0.168	0.154	0.163	0.257	0.243	0.205	0.178	0.164	0.18
F4	0.203	0.171	0.162	0.285	0.25	0.239	0.197	0.174	0.155	0.276	0.267	0.238	0.22	0.184	0.168
A5	0.226	0.266	0.283	0.161	0.169	0.191	0.225	0.25	0.277	0.146	0.159	0.184	0.226	0.262	0.279
B5	0.208	0.227	0.238	0.176	0.161	0.167	0.19	0.214	0.235	0.164	0.156	0.161	0.197	0.226	0.242
C5	0.191	0.202	0.224	0.201	0.17	0.156	0.172	0.194	0.22	0.204	0.172	0.152	0.166	0.19	0.217
D5	0.192	0.184	0.202	0.227	0.212	0.181	0.16	0.168	0.195	0.234	0.198	0.168	0.158	0.162	0.196
E5	0.201	0.19	0.196	0.27	0.225	0.204	0.178	0.162	0.167	0.26	0.235	0.2	0.167	0.153	0.165
F5	0.224	0.196	0.192	0.278	0.269	0.238	0.217	0.177	0.163	0.288	0.262	0.228	0.214	0.163	0.154

After removing outliers, to obtain an entry for a given bi-tap in the empirical bi-tap table representative of a generic user, one can no longer simply take the mean of the remaining data points. This would not give equal weight to each of the users. For example, if there is one outlier in the set of 25, belonging to, say, Participant 3, then if we define the empirical bi-tap table entry for that bi-tap as the mean of the remaining 24 points, Participant 3, having contributed only 4 points to the mean, will be underrepresented by a factor 4/5.

The empirical bi-tap table entries are instead calculated giving equal weight to each user: For a given bi-tap, take the mean time for each of the users on the points remaining in their quintuple after removing outliers, then define the empirical bi-tap table entry to be the average of those five means.

Formally, we calculate the entries as follows. First, we calculate individual bi-tap tables E_s for each test participant s, and then define as the mean of these tables:

$$\bar{E}(k_1, k_2) = \frac{\sum_{s \in S} E_s(k_1, k_2)}{|S|} \quad (2)$$

where k_1, k_2 are keys, S is the set of test participants, and $|S|$ is the cardinality of S (in our case, 5). The (k_1, k_2)th entry $E_s(k_1, k_2)$ of the individual bi-tap table E_s of test participant $s \in S$ is defined as the mean of the data points for s that remain for the bi-tap (k_1, k_2) after removing outliers using the procedure outlined earlier.

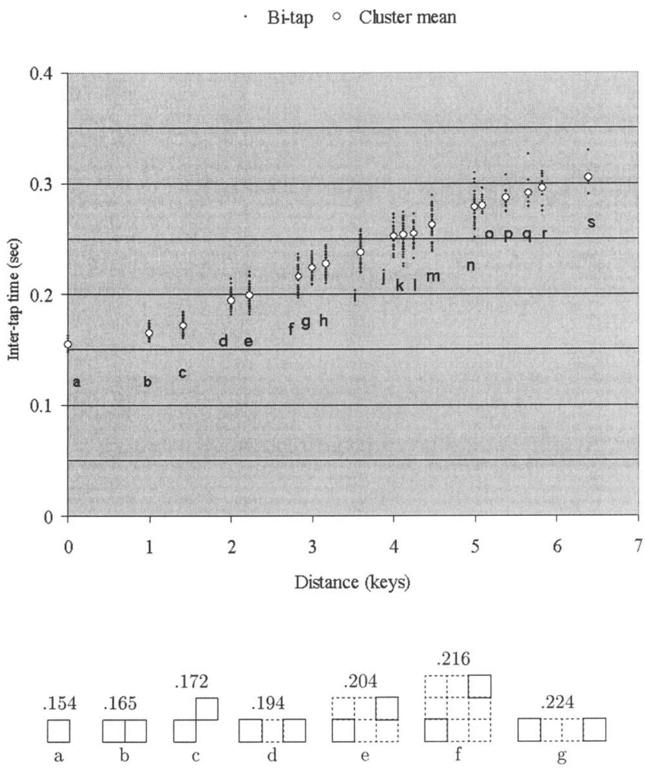
3.4. Interpreting the Empirical Data

Figure 5 is a scatterplot of bi-tap time against bi-tap length for the 900 bi-tap times in the empirical bi-tap table $\bar{E}$ (Figure 4). Distance is measured in key widths. As one would expect, bi-tap time increases with distance.

The vertical spread of the clusters (a cluster is the set of bi-tap times of a given length) is not random. Two patterns are lost in the projection of the data onto a time-distance scatter: Although the time interval between successive taps on a keyboard depends principally on the distance between them, it also depends on *position* and *trajectory.*

These dependencies are illustrated in Figure 6. The left grid shows the three top points of each of clusters *b*, *c*, and *d*, and the right grid shows the bottom three points of the clusters. The fast bi-taps are around the middle columns C and D, heading north/northeast/east; the slow bi-taps are around the left, bottom, and lower boundaries of the grid, heading west/southwest.

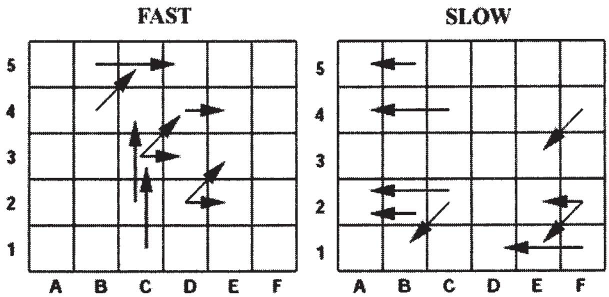
Figure 5. **(*Upper*) Time-distance scatter plot of the 900 bi-tap times in the empirical bi-tap table (Figure 4), the aggregated results from our experiments. (*Lower*) Chart illustrating a typical relation (modulo rotation and reflection) between keys of bi-taps in clusters *a* to *g* of the plot. For example, cluster *a* consists of double-taps on the same key (30 data points), cluster *b* consists of bi-taps between immediate neighbors (98 data points: 24⟨, 24®, 25◊, 25⇓), cluster *c* consists of bi-taps between diagonal neighbors (80 data points: 20™, 20⎛, 20©, 20∑), and so on. Cluster means are quoted above the corresponding pictures (.154, .165, ...).**



3.5. Relation With Fitts' Law

Fitts' law is a well-known model of human movement (Fitts, 1954; MacKenzie, 1991) that has been used in a number of papers on stylus keyboards (Hunter et al., 2000; Lewis et al., 1999; MacKenzie et al., 1994; MacKenzie & Zhang, 1999; MacKenzie et al., 1999; Soukoreff & MacKenzie, 1995) in the following form:

Figure 6. **Fast and slow bi-taps, showing how bi-tap time depends not only on the distance between source and target but also on position and trajectory. The left grid shows the top three points of each of clusters *b*, *c*, and *d* of Figure 5 (time-distance scatter plot), and the right grid shows the bottom three points of each of *b*, *c*, and *d*.**



$$MT_{ij} = b\log_2\left(\frac{A_{ij}}{W_j}+1\right) \tag{3}$$

where

MT_{ij} = mean time to move from key i to key j (in seconds)
W_j = size of key j
A_{ij} = distance from key i to key j
b = 1/4.9 is a fitted constant (MacKenzie, Sellen, & Buxton, 1991).

Fitts' law is somewhat inaccurate at a small scale (see, e.g., Section 2.2 of MacKenzie, 1991). This observation is confirmed by the data collected in our experiment: Figure 7 is the time-distance scatter of our empirical bi-tap table, with Fitts' law (in the aforementioned form) superimposed. Another difference between Fitts' law and our empirical results is the dependency of stylus dexterity on position and trajectory, as depicted in Figure 6. Fitts' law, as applied to a rectangular grid of square keys, is a translation and direction invariant.

3.6. Peak Input Rates of Stylus Keyboards

In this section, we use the approach described in Section 2 to predict peak expert input rates on various stylus keyboards. We begin with a simple illustrative example, an ABC Keyboard.

Define K_{ABC} to be the ABC keyboard layout depicted in Figure 8. (Note that keys D1, E1, and F1 are unused.) With the bi-gram probability table of the Appendix and the empirical bi-tap table (Figure 4), our model predicts the following peak expert text-input rate for the ABC layout:

Figure 7. **Relation between Fitts' law and our empirical data.**

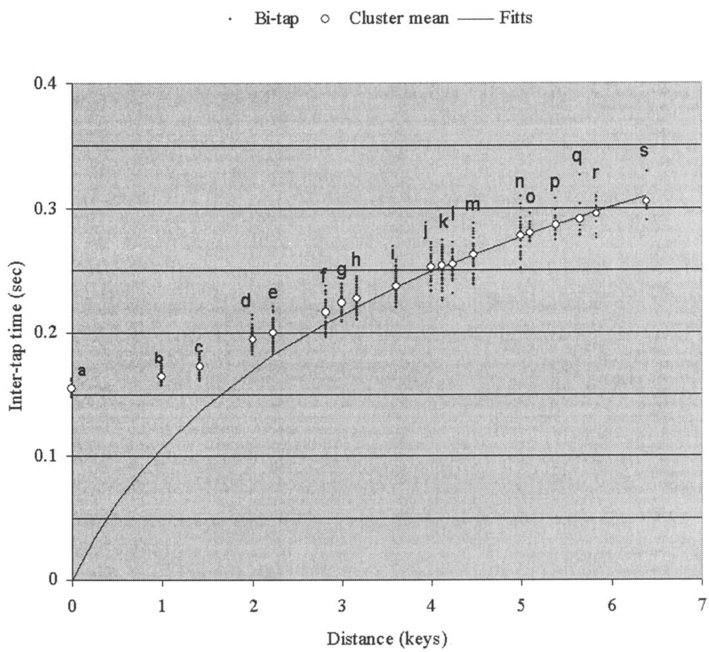


Figure 8. **Four keyboard layouts. Shaded keys are unused keys.**

Q	F	U	M	C	K	Z
SPACE		O	T	H	SPACE	
B	S	R	E	A	W	X
SPACE		I	N	D	SPACE	
J	P	V	G	L	Y	

OPTI

Z	V	C	H	W	K
F	I	T	A	L	Y
SPACE		N	E	SPACE	
G	D	O	R	S	B
Q	J	U	M	P	X

FITALY

A	B	C	D	E	F
G	H	I	J	K	L
M	N	SPACE		O	P
Q	R	S	T	U	V
	W	X	Y	Z	

ABC-center

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	SPACE			

ABC

$$R(K_{ABC}, \bar{B}, \bar{E}) = 4.702\text{char} / \text{sec} = 56.42\text{wpm}$$

where wpm denotes words per minute, with a word defined as 5 characters (including Spaces).

3.7. Validation of the Bi-Tap Table

In this section we compare our peak input rate predictions for two of the keyboard layouts shown in Figure 8 (the OPTI and the FITALY) with previous measures of peak input rate, thus validating the user model described by our empirical bi-tap table.

The OPTI Validation

In MacKenzie and Zhang (1999), users were trained over 20 sessions to tap 70 stock phrases on the OPTI layout depicted in Figure 8. Average text input rates followed the power law of learning from 17 wpm (1.42 char/sec) in Session 1 to 44.2 wpm (3.7 char/sec) in Session 20, and a regression ($R^2 = .997$) predicted a performance of 60.7 wpm (5.06 char/sec) on Session 50.

As can be seen in Figure 9, our predictions of peak text input rate correspond nicely with the asymptotic predictions of MacKenzie and Zhang. $\text{OPTI}_{\text{lower right}}$ denotes use of the OPTI with a fixed choice of Space key as the lower right of the four alternatives, and $\text{OPTI}_{\text{last}}$ denotes use of the OPTI where the user chooses the Space key closest to the last letter tapped. The details of the calculations, including the construction of the 5 × 7 bi-tap table, can be found in Section 3.1.

The FITALY Validation

The FITALY (Textware Solutions, 2000) is a commercially available stylus keyboard, the layout of which is shown in Figure 8. Figure 10 reproduces the results of a promotional competition held by the manufacturers in which contestants were timed tapping the following 181-character paragraph on the FITALY:

> What you need to do to have a chance to win the contest is to tap this sentence as fast as you can without any error. One more thing you need to have for a valid entry is a witness.

Our predictions of peak expert text input rate are interleaved in the table (Figure 10). Video recordings of human performance (available on the

Figure 9. **Validating our bi-tap table by comparing peak expert text input rate predictions with results of MacKenzie and Zhang (1999) on the OPTI keyboard depicted in Figure 8. See Section 3.11 for details of the calculations.**

	Characters/Second	Words Per Minute
MacKenzie and Zhang Session 50	5.06	60.7
$R(\text{OPTI}_{\text{lower right}}, \bar{B}, \bar{E}_7)$	5.10	61.2
$R(\text{OPTI}_{\text{last}}, \bar{B}_*, \bar{E}_7)$	5.25	63.0

Figure 10. **Validating our peak expert text input rate predictions using the commercial FITALY keyboard, the layout of which is depicted in Figure 8. The table is reproduced from the FITALY Web site (Textware Solutions, 2000), together with the interleaving of our predictions R_q (for q indicating various different patterns of Space key choice, detailed in Section 3.13) of peak expert text input rates. Shown are the performances of the top 10 competitors in a June to July 2000 speed-tapping competition. The competition task was to tap the 181-character paragraph quoted in Section 3.7.**

Competitor/Prediction	Characters/Second	Words Per Minute
1	6.165	73.98
2	5.803	69.64
3	5.798	69.47
4	5.607	67.38
R_{best}	**5.349**	**64.18**
R_{last}	**5.329**	**63.95**
5	5.288	63.46
R_{right}	**5.255**	**63.06**
R_{random}	**5.223**	**62.68**
6	5.191	62.29
R_{left}	**5.189**	**62.27**
7	5.130	61.56
8	4.805	57.66
9	4.756	57.07
10	4.694	56.33

Note. Char/sec = characters per second; wpm = words per minute.

FITALY Web site) are of individuals remarkably highly trained on entering this 181-character sequence, and as such represent peak input speeds. The fact that our predictions lie within the table adds to the validation of the bi-tap table for measuring peak input speed. Note that the sample test for the competition contains punctuation: two periods and two capitalizations. Hence, our predictions will be marginally too high. However, they still fall in essentially the same positions within Figure 10.

3.8. Stability With Respect to Bi-Gram Table

Our chosen bi-gram probability table is a composite of three bi-gram frequency tables B_1, B_2, and B_3. These three tables are reproduced in the Appendix. B_1 is Soukoreff and MacKenzie's (1995) extension of Mayzner and Tresselt's (1965) 26 × 26 table to include the Space character. B_2 is the 26 × 26 bi-gram table of Konheim's (1981) introductory cryptography textbook, to which we have added the same space bi-gram extension.[2] There are discrepancies between B_1 and B_2, possibly due to the fact that they were built from small text corpora. The following discrepancies[3] are the most notable: OF (80 vs. 731), ON (598 vs. 1232), TI (252 vs. 865), OU (1115 vs. 533), TH (3774 vs. 2161), HE (3155 vs. 2053).

To reduce these discrepancies we created a third bi-gram table B_3 of our own from a corpus 10 times the size, a mixture of informal and formal English (e-mail and classic novels). Stop-lists were used on proper nouns, and so forth. See Manning and Schütze (1999) for techniques for sampling data from text corpora. Then we defined $\bar{B}$ as the normalization of the weighted mean of (appropriate rescalings of) B_1, B_2, and B_3 (see Appendix). Figure 11 shows predictions of peak expert text input rate for the ABC keyboard with the four choices of bi-gram table. The observed rates did not vary significantly with selection of the bi-gram table.

3.9. Optimizing Stylus Keyboard Layout

The problem of finding the optimal layout for a stylus keyboard is equivalent to the problem of minimizing the average time between tapping two keys. Recall from Section 2 that the average time between tapping two keys on keyboard layout K with input text modeled by the bi-gram probability table P and stylus dexterity modeled by the bi-tap table E, is

$$\bar{M}(K,P,E) = \sum_{\alpha,\beta} P(\alpha,\beta)E(K(\alpha),K(\beta)) \tag{4}$$

where α and β range over the character set. Fixing $P = \bar{B}$ (see Appendix) and $E = \bar{E}$ (the empirical bi-tap table; Figure 4), our task is to minimize this expression with respect to K, a function from characters to keys. This has the

---

2. Because Konheim's total count of A to Z bi-grams (67,227) is nearly identical to that of Mayzner and Tresselt (67,320), we can conveniently add the Space bi-grams with only minor renormalization.

3. Note that direct comparisons of bi-gram frequencies between the tables makes sense because the sums of entries are nearly identical.

Figure 11. **Stability of peak expert text input rate prediction with respect to the bi-gram table *B*, with the ABC keyboard. $\bar{E}$ is the empirical bi-tap table shown in Figure 4.**

Bi-gram table B	$R(\text{ABC}, B, \bar{E})$	
	char/sec	wpm
$\bar{B}$	4.702	56.42
B_1	4.691	56.28
B_2	4.693	56.32
B_3	4.723	56.68

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	SPACE			

form of a standard optimization problem called the *quadratic assignment problem* (QAP).

The QAP (Koopmans & Beckman, 1957) has been shown to be an extremely hard problem. Not only is it NP-hard (Sahni & Gonzalez, 1976), but it is NP-hard to approximate its optimal solution to within any constant factor (Queyranne, 1986). There are a number of heuristics that can be employed to find reasonable solutions to the QAP: genetic algorithms, the Metropolis method, and dynamic simulation, to name but a few. We chose to implement the hybrid ant system proposed by Gambardella, Taillard, and Dorigo (1997) because it has been shown to find quality solutions quickly.

The best solution found by the hybrid ant system was the keyboard layout $K_{\bar{B},\bar{E}}$ depicted in Figure 12, with a predicted peak expert text input rate of

$$R(K_{\bar{B},\bar{E}}, \bar{B}, \bar{E}) = 5.438\,\text{char/sec} = 65.26\ \text{wpm}.$$

See Section 3.11 for a comparison with other keyboards.

3.10. Variation of Best Solution With Respect to Bi-Gram Table

Figure 13 shows how the best solution produced by the hybrid ant system varies with respect to the bi-gram table parameterizing the optimization problem. One can observe how the final layout is directly related to the idiosyncrasies of a particular bi-gram table. For example, recall the major discrepancies between the frequency tables B_1 (Mayzner and Tresselt) and B_2 (Konheim): OF (80 vs.731), ON (598 vs. 1232), TI (252 vs. 865), OU (1115 vs. 533), TH (3774 vs. 2161), HE (3155 vs. 2053). Notice how the strong preference of B_2 for

Figure 12. **The keyboard layout $K_{\bar{B},\bar{E}}$, the best solution found by the hybrid ant system to the quadratic assignment problem of optimal stylus keyboard layout. The predicted peak expert text input rate $R(K_{\bar{B},\bar{E}}, \bar{B}, \bar{E})$ is 5.438 char/sec = 65.26 wpm.**

	K	G	I	C	Z
	F	N	T	H	W
Q	O	S	SPACE	A	Y
J	U	R	E	D	V
	P	M	L	B	X

OF, ON, and TI over B_7 (by an order of magnitude in the case of OF) is observable in the layouts: the keys of the pairs OF and IT are directly adjacent, and the keys O and D of ON are diagonally adjacent. These three bi-grams stretch over distances of 3.2, 2, and 3.2 key widths, respectively. Figure 14 shows the variation in peak input speed of the four optimized layouts under evaluation with respect to each of the four bi-gram tables.

3.11. Benchmarking Various Keyboards

Using our predictions of peak expert text input rate, we benchmark four keyboard layouts against the simple ABC layout: the OPTI, the FITALY, the best solution $K_{\bar{B},\bar{E}}$ to the keyboard layout optimization problem discovered by the hybrid ant system, and a variant of the ABC with Space at the center, which we call ABC-center. The five layouts in question are depicted in Figures 8 and 12.

The results are shown in Figure 15. Observe that simply moving the Space key to the center of the ABC already increases performance by nearly 4%. The FITALY yields an additional 9.4% increase relative to the ABC. Our ant algorithm solution $K_{\bar{B},\bar{E}}$ gains yet another 2.3% above the FITALY. These results are not definitive because a sample size of only 5 users was used to generate our table of bi-tap data.

In the following sections we report in detail the calculations involved in doing the simulations for the FITALY and OPTI. They were nontrivial because the layouts have multiple Space keys.

3.12. Predicting the Peak Expert Text Input Rate of the OPTI

Concerning their OPTI layout, MacKenzie and Zhang (1999) noted that "having four SPACE keys is convenient; but, using the optimal SPACE key re-

Figure 13. **Variation of best solution found by the hybrid ant system optimization with respect to the bi-gram probability table.**

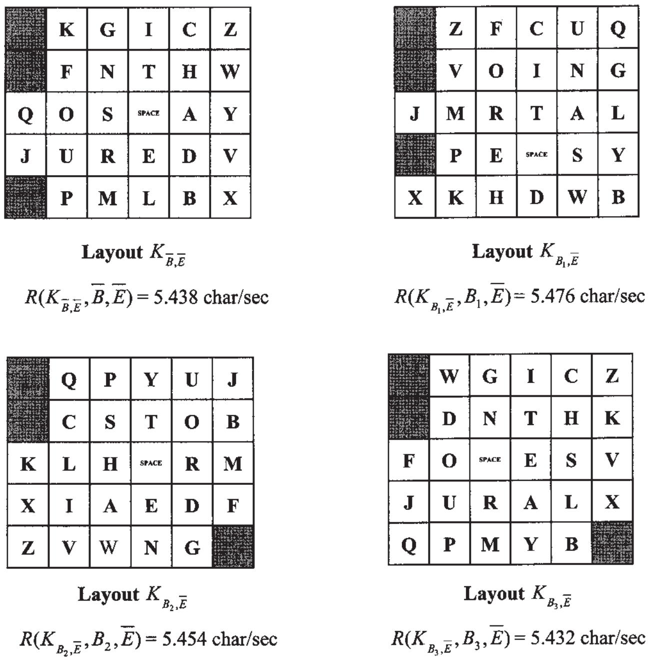


Figure 14. **Stability of the peak expert text input rate with respect to choice of bi-gram table. Entries in the table are in words per minute. The row label is the bi-gram table used in the ant algorithm search; the column label indicates the bi-gram table used in the peak expert text input rate prediction *R*. Thus each row consists of four different evaluations of a keyboard.**

Bi-gram Table Used for Ant Algorithm	Bi-gram Table Used for Prediction Rates			
	B1	B2	B3	$\bar{B}$
B1	65.486 wpm	64.709 wpm	64.707 wpm	64.935 wpm
B2	64.650 wpm	65.443 wpm	64.643 wpm	64.869 wpm
B3	65.112 wpm	65.123 wpm	65.279 wpm	65.137 wpm
$\bar{B}$	65.265 wpm	65.345 wpm	65.202 wpm	65.242 wpm

Note. wpm = words per minute.

quires extra judgment on-the-fly and this is not likely to occur—at least within the confines of the limited practice in this study" (p. XXX).

Hence, for our predictions, we considered the two patterns of Space use observed by MacKenzie and Zhang in the penultimate section of their paper: (a)

Figure 15. **Benchmarking results for the five keyboard layouts depicted in Figures 8 and 12. The right column shows the percentage performance increase above the ABC. The subscripts on the FITALY and OPTI indicate the pattern of space usage (see Section 3.7 for details). Because on-the-fly calculations of optimal tri-character path are somewhat unrealistic to expect from users (see MacKenzie & Zhang, 1999), we suggest $\text{FITALY}_{\text{last}}$ as the most representative of FITALY performance. Predictions of peak expert input speeds are not estimates of real continuous performance (see Section 2.1).**

	Predicted Peak Expert Text Input Rate		
Keyboard K	**Char/sec**	**wpm**	**% > ABC**
$K_{\bar{B},\bar{E}}$	5.438	65.26	15.65
$\text{FITALY}_{\text{best}}$	5.349	64.19	13.77
$\text{FITALY}_{\text{last}}$	5.329	63.95	13.35
$\text{FITALY}_{\text{right}}$	5.255	63.06	11.77
$\text{OPTI}_{\text{last}}$	5.248	62.98	11.63
$\text{FITALY}_{\text{random}}$	5.223	62.67	11.08
$\text{FITALY}_{\text{left}}$	5.189	62.27	10.37
$\text{OPTI}_{\text{lower right}}$	5.103	61.24	8.54
ABC-center	4.888	58.65	3.95
ABC	4.702	56.42	0.00

Note. char/sec = characters per second; wpm = words per minute.

the closest Space key to the last character tapped, denoted $\text{OPTI}_{\text{last}}$; and (b) the lower right Space key (of the four, chosen because right-handers may be naturally inclined to pick the Space key that obscures their view of the grid the least), denoted $\text{OPTI}_{\text{lower right}}$.

Our empirical bi-tap table $\bar{E}$ (Figure 4) was obtained for a 5 × 6 grid, but the OPTI has a 5 × 7 grid. Rather than repeating our experiments in full with a 5 × 7 grid, we constructed a 5 × 7 empirical bi-tap table $\bar{E}_7$ from $\bar{E}$ as follows:

- The bi-tap values in the left 5 × 6 subgrid of $\bar{E}_7$ are taken directly from $\bar{E}$.
- Making the simplifying assumption that the inner half of each Space key is used, this leaves only two keys unaccounted for in column 7—Z and X.
- The bi-gram probabilities between characters of column 1 (Q, B, J) and column 7 (Z, X) are all zero; hence, one can assign arbitrary times to the bi-taps of $\bar{E}_7$ between column 1 and column 7 (we chose 0 sec) without any effect on the model.
- This leaves the bi-tap times between Z and X and columns 2–7. Because this is a 5 × 6 subgrid, we used $\bar{E}$ (shifted right by one column).

In essence, we are thinking of the OPTI as the 5 × 6 grid consisting of columns 1–6, with the infrequently used X and Z "pasted" on the side. Note that the use of this artificially constructed 5 × 7 table in peak expert input speed prediction should be reasonably accurate because the only bi-grams that are not covered by the 5 × 6 case are the 0 probability pairings between {Q, B, J} and {Z, X}.

OPTI Prediction: Lower Right Space Choice

The prediction of peak expert text input rate $R(\mathrm{OPTI}_{\text{lower-right}}, \bar{B}, \bar{E}_7)$ in Figure 9 was based on the following data. The character set was C_{27} = {A–Z, Space}. The bi-gram probability table $\bar{B}$ is from the Appendix, and the bi-tap table $\bar{E}_7$ is that derived from $\bar{E}$ (Figure 4) in the manner described earlier.

OPTI Prediction: Closest-to-Last Space Choice

The prediction of peak expert text input rate $R(\mathrm{OPTI}_{\text{last}}, \bar{B}, \bar{E}_7)$ in Figure 9 was based on the following data. The keyboard function $\mathrm{OPTI}_{\text{last}}$, assigning keys to characters, was the same as $\mathrm{OPTI}_{\text{lower right}}$ on characters A to Z, and with the following assignment of Space characters (lower left, lower right, upper left, upper right) to keys:

$\mathrm{SPACE}_{\text{ll}}$	$\mathrm{SPACE}_{\text{lr}}$	$\mathrm{SPACE}_{\text{ul}}$	$\mathrm{SPACE}_{\text{ur}}$
B2	F2	B4	F4

where we use the key-naming convention of Figure 3 with an additional right-hand column G1 to G5. The bi-tap table $\bar{E}_7$ is that derived from $\bar{E}$ in the manner described at the beginning of Section 3.12.

The bi-gram probability table $\bar{B}_*$ used is $\bar{B}$ on A to Z, together with bi-grams involving the four Space characters calculated as follows:

- The probability of consecutive spaces is zero.
- Given an A to Z character a and one of the four spaces σ, the probability (α, σ) of the bi-gram (α, σ) is (α, Space) if the σ key is the closest space to the α key on the grid (picking right-most and lowest in case of tie-break), and 0 otherwise.
- Given an A to Z character α and one of the four spaces σ, the probability $\bar{B}_*$ of the bi-gram (α, σ) is $\theta_\sigma\, \bar{B}$(Space, α), where θ_σ is the proportion of the time σ occurs as a trailing space:

$$\theta_\sigma = \frac{\sum_{\beta\in\{A,\ldots,Z\}} \bar{B}_*(\beta,\sigma)}{\sum_{\beta\in\{A,\ldots,Z\}} \sum_{X=ll,lr,ul,ur} \bar{B}_*(\beta,\text{Space}_X)} \tag{5}$$

3.13. Predicting the Peak Expert Text Input Rate of the FITALY

The cases $\text{FITALY}_{\text{last}}$, $\text{FITALY}_{\text{right}}$, and $\text{FITALY}_{\text{left}}$ in Figure 15 are analogous to $\text{OPTI}_{\text{last}}$, $\text{OPTI}_{\text{lower right}}$, and $\text{OPTI}_{\text{lower left}}$, respectively. $\text{FITALY}_{\text{random}}$ denotes a random choice of Space key: We used the bi-gram table derived from that splits the probability of bi-grams of $\bar{B}$ involving Space equally in two.

The case $\text{FITALY}_{\text{best}}$ was labor intensive. We considered by hand all 676-character tri-grams of the form (α, Space, β) for alphabet characters α and β, and decided which of the two Space keys would be the optimal choice. Having assigned left/right to each such tri-gram, the probability of the bi-gram (α, $\text{Space}_{\text{left}}$) is

$$\bar{B}(\alpha,\text{Space}) = \frac{\sum_{B\in L(\alpha,\beta)} \bar{B}(\text{Space},\beta)}{\sum_{\beta} \bar{B}(\text{Space},\beta)} \tag{6}$$

where β ranges over the character set, and $L(\alpha, \beta)$ is the subset of characters for which the tri-gram (α, Space, β) was designated as using the left Space key. The other space bi-gram probabilities work similarly.

Note that, in the absence of tri-gram data, one has to assume a uniform distribution of tri-grams with respect to bi-grams. The effect of this assumption should be negligible.

3.14. Stylus Keyboards: Conclusion

We illustrated the empirical bi-action table technique in the context of stylus keyboard design. A hybrid ant system yielded an optimized stylus keyboard layout (Figure 12) that outperformed the commercial FITALY layout. This result, however, is not to be considered final due to the small number of 5 participants used to generate our table of bi-tap data and the coincidence of the training data and test data.

There is an observation that stems from our experiments: All previous related work on stylus keyboard design has only considered distance between two keys as a predictor of the duration of the motion between these keys. We found that this duration depends also on the first key position and on the relative position of the second key (see Figure 6). These dependencies, and any other more subtle dependencies that could be impossible to model, are automatically taken into account by the empirical bi-action table.

Within the realm of stylus keyboard design, the following topics are possibilities for future work:

- *Bi-tap corpus:* Although 25 data points for each of 900 bi-taps is a considerable amount of data, a clear route to improving our work would be to obtain a much larger corpus of test data. Just as bi-gram tables based on large corpora of text are publicly available to those who which to use it (e.g., in cryptology or natural language processing), it would make sense to make publicly available a bi-tap table built on a considerable corpus of stylus dexterity test data. Researchers interested in benchmarking and optimizing their own layouts (e.g., with a variety of different character sets, perhaps from a variety of languages) could make use of the corpus bi-tap table.
- *Errors:* It would be nice to incorporate a quantitative account of errors.
- *Obscuration:* As a right-hander, having just tapped a key in the top left of the grid, when aiming for a key in the mid- or lower right, is there any delay due to the fact that my hand is obscuring the target zone?
- *Investigate position dependency:* A possible explanation for the dependency of bi-tap time on position (preference for the center; see Figure 6) is that the users rest on the side of the PDA with the wrist, outer-palm, and/or outer edge of little finger. The natural resting position of the stylus is somewhere over the center of the grid. Motions around the edges of the grid require either a cramped or overstretched position of the fingers and thumb, or a cocking of the wrist.
- *Investigate trajectory dependency:* A possible explanation for the dependency of bi-tap times on trajectory (dislike for heading west; see Figure 6) is that everyday handwriting is from left to right. Low-level motor skills in the hand involved in moving against the direction of writing are probably less well developed. To further investigate trajectory dependency one could carry out the tests with left-handers and/or people whose mother tongue is written from right to left.
- *Optimization incorporating double-Space and/or double-E:* The problem of finding the optimal keyboard with two Space keys and/or two E keys is much harder than the quadratic assignment problem. It would be useful to determine whether having two Space keys can yield a faster layout.
- *Tri-taps:* Although the extent of dependency of bi-taps on the preceding tap is probably low, it may be useful to collect data for an empirical tri-tap table. However, with 27,000 possible triples (in the case of 30 keys), experiments would be impractical.

- *Keyboard shape:* Hunter et al. (2000) and Zhai et al. (2000) considered layouts on a hexagonal grid. One could also imagine a radial pattern, a dartboard shape with Space as the "bull's-eye." Empirical bi-tap tables could be obtained for these shapes and used to produce optimized key layouts. One observation of a rectangular grid arrangement is that it maximizes the number of adjacent keys. Define a *neighbor* of a key to be any other key that is reachable without having to leap over an intermediate key. In a hexagonal grid, each key has only six neighbors, whereas in a rectangular grid each key has eight. Our empirical data (Figure 5) show an observable jump between neighbor bi-tap times (clusters *a*, *b*, and *c*) and times of bi-taps involving leaps (cluster *d* and beyond). This neighborhood property may be particularly important with regard to the number of neighbors of the central Space key, by far the most frequently used key, in which case a rectangular arrangement might confer an advantage over a hexagonal pattern.
- *x/y-Scale:* The commercial version of the FITALY has rectangular keys that are longer in the horizontal direction. Does such a feature speed up or slow down text input?

4. CONCLUSION

In this article we presented a technique for predicting peak expert input speeds on text input systems. The technique is intended as a tool in the interface development cycle between initial evaluations using abstract mathematical models (e.g., Fitts' law, Hick's law, and the power law) and final evaluations by full empirical testing. We illustrated the approach in the context of stylus keyboards. Empirical bi-action tables could be used in the analysis and design of a wide variety of text input systems, such as two-handed keyboards, chording keyboards, cell phones, glove gesture input, and forms of stylus input including Graffiti, Quikwriting (Perlin, 1998), and Unistrokes (Goldberg & Richardson, 1993).

Relative to full empirical testing, the independence of the empirical bi-action table E from the logical aspect of the input system (the character map) confers the following advantages: (a) A change in character mapping (e.g., a change in a keyboard layout) does not demand fresh experimental trials; (b) we avoid the cost of training participants up to expert level with a particular character mapping (e.g., keyboard layout); and (c) having obtained E, we can perform an algorithmic search for optimal character mappings. One disadvantage is the reduction in accuracy due to the higher level of abstraction.

Relative to mathematical modeling with laws we cite two advantages: (a) greater generality in the sense of coverage of the full range of input systems, including those for which laws do not easily apply (e.g., glove gestures) and (b) greater accuracy due to the higher specificity of empirical testing. Disadvantages include the cost of empirical testing and the necessity of undertaking new tests for each new input system, aside from the case of variations in character map.

We have not discussed issues such as ease of use or repetitive strain injury. The appropriate balance between such issues and the optimization of input speed should be borne in mind by any researcher choosing to employ our technique.

NOTES

Acknowledgments. We acknowledge Julien Basch, Vaughan Pratt, the participants of our experiments, and funding from the Stanford Wearable Computing Laboratory. We are appreciative for extremely insightful feedback from Scott MacKenzie and three anonymous referees.

Authors' Present Addresses. Dominic Hughes, Gates Building, Room 486, Computer Science Department, Stanford University, Stanford, CA 94305. E-mail: dominic@cs.stanford.edu. James Warren, Gates Building, Room 118, Computer Science Department, Stanford University, Stanford, CA 94305. E-mail: warren@sccm.stanford.edu. Orkut Buyukkokten, Gates Building, Room 430, Computer Science Department, Stanford University, Stanford, CA 94305. E-mail: orkut@stanford.edu.

HCI Editorial Record. First manuscript received November 6, 2001. Accepted by I. Scott Mackenzie. Final manuscript received May 29, 2001. — *Editor*

REFERENCES

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

Fitts, P. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology, 47,* 381–391.

Gambardella, L., Taillard, D., & Dorigo, M. (1999). Ant colonies for the QAP. *Journal of the Operational Research Society, 50,* 167–176.

Goldberg, D., & Richardson, C. (1993). Touch-typing with a stylus. *Proceedings of INTERCHI 93 Conference on Human Factors in Computing Systems*. New York: ACM.

Hunter, M., Zhai, S., & Smith, B. (2000). Physics-based graphical keyboard design (interactive poster). *Proceedings of the CHI 2000 Conference on Human Factors in Computing Systems.*

Konheim, A. (1981). *Cryptography: A primer.* New York: Wiley.

Koopmans, T., & Beckman, M. (1957). Assignment problems and the location of economic activities. *Econometrica, 25,* 53–76.

Lewis, J., LaLomia, M., & Kennedy, P. (1999). Evaluation of typing key layouts for stylus input. *Proceedings of the Human Factors and Ergonomics Society 43rd Annual Meeting.*

MacKenzie, I. (1991). *Fitts' law as a performance model in human–computer interaction.* Unpublished doctoral dissertation, University of Toronto, Canada.

MacKenzie, I., Nonnecke, R., McQueen, J., Riddersma, S., & Meltz, M. (1994). A comparison of three methods of character entry on pen-based computers. *Proceedings of the Human Factors and Ergonomics Society 38th Annual Meeting (Human Factors and Ergonomics Society).*

MacKenzie, I., Sellen, A., & Buxton, W. (1991). A comparison of input devices in elemental pointing and dragging tasks. *Proceedings of the CHI 99 Conference on Human Factors in Computing Systems.*

MacKenzie, I., & Zhang, S. (1999). The design and evaluation of a high performance soft keyboard. *Proceedings of the CHI 99 Conference on Human Factors in Computing Systems.*

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

Manning, C., & Schütze, H. (1999). *Statistical natural language processing.* Cambridge, MA: MIT Press.

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

Noel, R., & McDonald, J. (1989). Automating the search for good designs: About the use of simulated annealing and user models. *Proceedings of Interface 89 (Human Factors Society).*

Perlin, K. (1998). Quikwriting: Continuous stylus-based text entry. *Proceedings of the UIST 98 Symposium on User Interface Software and Technology.*

Queyranne, M. (1996). Performance ratio of heuristics for triangle inequality quadratic assignment problems. *Operations Research Letters, 4,* 231–234.

Sahni, S., & Gonzalez, T. (1976). P-complete approximation problems. *Journal of the ACM, 23,* 555–565.

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

Textware Solutions. (2000). The FITALY one-finger keyboard. Retrieved DATE from http://www.twsolutions.com/domperignon/domperignon2.htm

Zhai, S., Hunter, M., & Smith, B. (2000). The Metropolis keyboard: An exploration of quantitative techniques for virtual keyboard design. *Proceedings of the UIST 2000 Symposium on User Interface Software and Technology.*

APPENDIX

Bi-gram Frequency Tables

The bi-gram probability table used for peak expert text-input speed predictions is the normalization of the bi-gram frequency table depicted in Figure

A–1. The table in Figure A–1 is the equally weighted mean of the (appropriate rescalings of) bi-gram frequency tables B_1, B_2, and B_3 shown in Figures A–2, A–3, and A–4. The composite table in Figure A–1 was constructed to soften the idiosyncrasies of B_1 and B_2, as described in Section 3.8.

Figure A–1. **The composite bi-gram frequency table whose normalization *B* was used for predicting peak expert text input rates.**

1st/2nd	a	b	c	d	e	f	g	h	i	j	k	l	m	n
a	288	13549	25757	31409	481	5729	12790	815	29360	757	9581	60550	17202	127606
b	10479	899	108	60	40457	12	25	4	6509	1024	0	12961	477	12
c	31031	0	3225	48	35807	42	8	35759	11285	0	12253	9831	132	124
d	11842	183	137	3508	44599	290	1970	52	20818	254	42	2398	1023	847
e	51653	1986	22422	77414	36532	10572	5980	1478	10417	214	2617	32996	20979	87818
f	11893	4	0	17	13550	8257	8	4	18093	0	0	4425	128	52
g	9145	31	25	87	25779	655	5853	19216	5766	0	714	3234	1764	5721
h	89735	155	163	110	232576	38	0	93	62045	0	0	525	392	639
i	11684	3535	36606	25487	19893	7693	18733	39	405	29	4834	29231	21402	137417
j	804	0	0	34	2427	0	0	0	435	0	0	0	0	0
k	710	35	0	43	23129	150	55	230	7929	8	124	783	83	5011
l	31972	407	483	22603	53784	4386	1761	17	35748	0	2604	44528	1842	100
m	35449	5107	232	17	52729	299	4	4	19136	4	4	275	6301	615
n	15884	200	19972	92846	49708	2879	68077	439	19725	598	5391	5501	1592	6121
o	4717	5779	7743	11304	2484	46891	3506	751	7393	405	7666	19636	39612	93263
p	17751	25	79	68	27376	146	29	3311	7246	0	29	14601	540	31
q	0	0	0	0	0	0	0	0	93	0	0	0	0	0
r	35016	1147	4275	13991	114733	1256	5414	1365	38598	51	6719	5794	9389	10996
s	18535	578	6451	505	59388	721	112	24381	29692	59	3867	4172	3570	1115
t	29959	149	2622	70	68067	261	97	262395	56590	0	12	7708	1134	546
u	6068	6778	10510	4137	7637	1031	10742	66	6028	0	157	24086	7017	27413
v	5133	0	25	62	60579	0	0	0	14432	0	0	8	39	0
w	38422	141	21	383	26190	17	0	31920	30218	0	56	852	12	7133
x	1672	0	1274	0	1208	21	4	52	1705	0	0	8	35	0
y	1023	512	220	189	9478	38	35	48	2527	0	31	628	773	364
z	822	0	0	0	2748	0	0	0	773	0	0	232	0	0
SPC	184447	89025	77608	48894	39111	88609	36904	120949	52582	7472	11872	58899	80066	42462

1st\2nd	o	p	q	r	s	t	u	v	w	x	y	z	SPC
a	288	11944	12	70150	63881	91864	7851	19696	4753	1414	23499	1060	20307
b	15192	21	0	7893	2528	966	15334	344	35	0	7624	0	2599
c	42145	0	147	8308	1281	19584	7075	0	0	0	1618	46	4105
d	14685	149	31	6377	7865	185	6373	1111	325	34	4183	0	219394
e	3475	10538	2014	132134	71565	31389	1465	16112	10289	10305	11296	278	415298
f	35100	8	0	15032	294	7001	5807	0	8	4	569	0	29183
g	7669	0	4660	6981	3095	2192	3220	0	31	186	300	4	59251
h	36088	62	0	6456	772	12546	4717	0	398	0	2287	0	59925
i	32513	4531	541	20973	65647	65916	504	14844	0	1391	35	2950	406
j	3205	0	0	31	0	0	4668	0	0	0	4	0	0
k	370	31	0	97	4031	56	149	0	55	0	530	0	24673
l	25575	1986	0	962	9276	6211	7119	1471	1003	0	25979	4	53418
m	23368	11325	4	159	4931	114	7520	31	8	0	7298	0	36885
n	31639	236	322	383	23368	54107	4515	1929	434	157	8192	87	113950
o	20460	7962	0	23236	3704	4174	6249	0	127	0	801	0	11740
q	0	0	0	0	0	0	7923	0	17	0	0	0	17
r	46514	2257	8	8762	26177	20340	8180	3747	970	31	15545	0	122685
s	22918	10336	593	334	22186	70788	16837	12	2250	0	2862	0	197773
t	60529	126	0	22881	19760	14855	12275	101	6798	0	10878	145	205492
u	451	9696	43	35360	27068	35532	199	114	72	169	425	188	22756
v	3132	0	0	4	66	0	96	31	0	0	625	0	97
w	19414	12	0	2420	2822	203	79	0	21	0	131	0	25815
x	79	3132	4	4	0	2371	261	0	17	163	46	0	1944
y	22990	1088	0	296	5802	707	189	0	270	0	62	48	104700
z	207	0	0	0	93	0	82	0	0	8	110	448	235
SPC	84385	52128	4297	42579	139178	330461	16716	11589	150110	106	37956	192	0

Figure A–2. **Bi-gram frequency table B_1, Soukoreff and MacKenzie's (1995) extension of Mayzner and Tresselt's (1965) 26 × 26 table to include the Space character. From "Theoretical upper and lower bounds on typing speed using a stylus and soft keyboard," by R. Soukoreff and I. MacKenzie, 1995, *Behaviour & Information Technology, 14,* p. XXX. Copyright 1995 by Taylor & Francis, Ltd. Reprinted with permission.**

1st/2nd	a	b	c	d	e	f	g	h	i	j	k	l	m	n
a	2	144	308	382	1	67	138	9	322	7	146	664	177	1576
b	136	14	0	0	415	0	0	0	78	18	0	98	1	0
c	368	0	13	0	285	0	0	412	67	0	178	108	0	1
d	106	1	0	37	375	3	19	0	148	1	0	22	1	2
e	670	8	181	767	470	103	46	15	127	1	35	332	187	799
f	145	0	0	0	154	86	0	0	205	0	0	69	3	0
g	94	1	0	0	289	0	19	288	96	0	0	55	1	31
h	1164	0	0	0	3155	0	0	1	824	0	0	5	1	0
i	23	7	304	260	189	56	233	0	1	0	86	324	255	1110
j	2	0	0	0	31	0	0	0	9	0	0	0	0	0
k	2	0	0	0	337	0	0	0	127	0	0	10	1	82
l	332	4	6	289	591	59	7	0	390	0	38	546	30	1
m	394	50	0	0	530	6	0	0	165	0	0	4	28	4
n	100	2	98	1213	512	5	771	5	135	8	63	80	0	54
o	65	67	61	119	34	80	9	1	88	3	123	218	417	598
p	142	0	1	0	280	1	0	24	97	0	0	169	0	0
q	0	0	0	0	0	0	0	0	0	0	0	0	0	0
r	289	10	22	133	1139	13	59	21	309	0	53	71	65	106
s	196	9	47	0	626	0	1	328	214	0	57	48	31	16
t	259	2	31	1	583	1	2	3774	252	0	0	75	1	2
u	45	53	114	48	71	10	148	0	65	0	0	247	87	278
v	27	0	0	0	683	0	0	0	109	0	0	0	0	0
w	595	3	0	6	285	0	0	472	374	0	1	12	0	103
x	17	0	9	0	9	0	0	0	10	0	0	0	0	0
y	11	10	0	0	152	0	1	1	32	0	0	7	1	0
z	3	0	0	0	26	0	0	0	2	0	0	4	0	0
SPACE	1882	1033	864	515	423	1059	453	1388	237	93	152	717	876	478

1st/2nd	o	p	q	r	s	t	u	v	w	x	y	z	SPACE
a	1	100	0	802	683	785	87	233	57	14	319	12	50
b	240	0	0	88	15	7	256	1	1	0	13	0	36
c	298	0	1	71	7	154	34	0	0	0	9	0	47
d	137	0	0	83	95	3	52	5	2	0	51	0	2627
e	44	90	9	1314	630	316	8	172	106	87	189	2	4904
f	429	0	0	188	4	102	62	0	0	0	4	0	110
g	135	0	0	98	42	6	57	0	1	0	2	0	686
h	487	2	0	91	8	165	75	0	8	0	32	0	715
i	88	42	2	272	484	558	5	165	0	15	0	18	4
j	41	0	0	0	0	0	56	0	0	0	0	0	0
k	3	1	0	0	50	0	3	0	0	0	8	0	309
l	344	34	0	11	121	74	81	17	19	0	276	0	630
m	289	77	0	0	53	2	85	0	0	0	19	0	454
n	349	0	3	2	148	378	49	3	2	2	115	0	1152
o	336	138	0	812	195	415	1115	136	398	2	47	5	294
p	149	64	0	110	48	40	68	0	3	0	14	0	127
q	0	0	0	0	0	0	66	0	0	0	0	0	0
r	504	9	0	69	318	190	89	22	5	0	145	0	1483
s	213	107	8	0	168	754	175	0	32	0	34	0	2228
t	331	0	0	187	209	154	132	0	84	0	121	1	2343
u	3	49	1	402	299	492	0	0	0	1	7	3	255
v	33	0	0	0	0	0	1	0	0	0	11	0	0
w	264	0	0	35	21	4	2	0	0	0	0	0	326
x	1	22	0	0	0	23	8	0	0	0	0	0	21
y	339	16	0	0	81	2	1	0	2	0	0	0	1171
z	2	0	0	0	3	0	0	0	0	0	3	9	2
SPACE	721	588	42	494	1596	3912	134	116	1787	0	436	2	0

Figure A–3. **Bi-gram frequency table B_2, from Konheim's introductory cryptography textbook, to which we have added the same space bi-gram extensions of Figure A2. Because Konheim's total count of A to Z bi-grams (67,227) is nearly identical to that of Mayzner and Tresselt (67,320), we can conveniently add the space bi-grams with only minor renormalization.**

1st/2nd	a	b	c	d	e	f	g	h	i	j	k	l	m	n
a	7	125	251	304	13	65	151	13	311	13	67	681	182	1216
b	114	7	2	1	394	0	0	0	74	7	0	152	6	0
c	319	0	52	1	453	0	0	339	202	0	86	98	4	3
d	158	3	4	33	572	1	20	1	273	5	0	19	27	8
e	492	27	323	890	326	106	93	16	118	4	27	340	253	1029
f	98	0	0	0	150	108	0	0	188	0	0	35	1	1
g	122	0	0	2	271	20	145	95	0	0	23	3	51	129
h	646	2	5	3	2053	0	0	2	426	0	0	6	6	14
i	236	51	476	285	271	80	174	1	10	0	31	352	184	1550
j	18	0	0	0	26	0	0	0	5	0	0	0	0	0
k	14	1	0	1	187	1	0	7	56	0	4	7	1	20
l	359	5	6	197	513	28	29	0	407	0	21	378	22	1
m	351	65	5	0	573	2	0	0	259	0	0	2	126	8
n	249	2	281	761	549	46	630	6	301	4	30	33	47	88
o	48	57	91	130	21	731	46	14	52	8	44	234	397	1232
p	241	0	1	0	310	0	0	42	75	0	0	144	13	1
q	0	0	0	0	0	0	0	0	0	0	0	0	0	0
r	470	15	79	129	1280	14	80	8	541	0	94	75	139	149
s	200	4	94	9	595	8	0	186	390	0	30	48	37	7
t	381	2	22	1	872	4	1	2161	865	0	0	62	27	9
u	72	87	103	51	91	11	80	2	54	0	3	230	69	318
v	65	0	0	2	522	0	0	0	223	0	0	0	1	0
w	282	1	0	4	239	0	0	175	259	0	0	5	0	44
x	9	0	15	0	17	0	0	1	15	0	0	0	1	0
y	17	1	3	2	84	0	0	0	20	0	1	5	11	5
z	18	0	0	0	36	0	0	0	17	0	0	1	0	0
SPACE	1882	1033	864	515	423	1059	453	1388	237	93	152	717	876	478

1st/2nd	o	p	q	r	s	t	u	v	w	x	y	z	SPACE
a	5	144	0	764	648	1019	89	137	37	17	202	15	50
b	118	0	0	81	28	6	89	2	0	0	143	0	36
c	606	0	1	113	23	237	92	0	0	0	25	0	47
d	111	0	1	49	75	2	91	15	6	0	40	0	2627
e	30	143	25	1436	917	301	36	160	153	113	90	3	4904
f	326	0	0	142	3	58	54	0	0	0	5	0	110
g	0	0	150	29	28	58	0	0	0	6	0	0	686
h	287	0	0	56	10	85	31	0	4	0	15	0	715
i	554	62	5	212	741	704	7	155	0	14	1	49	4
j	45	0	0	1	0	0	48	0	0	0	0	0	0
k	7	0	0	3	39	1	1	0	0	0	4	0	309
l	208	11	0	9	104	68	72	15	3	0	219	0	630
m	240	139	0	5	47	1	65	1	0	0	37	0	454
n	239	2	3	5	340	743	56	31	8	1	71	2	1152
o	125	164	0	861	201	223	533	188	194	7	23	2	294
p	268	103	0	409	32	51	81	0	0	0	3	0	127
q	0	0	0	0	0	0	73	0	0	0	0	0	0
r	510	25	0	97	300	273	88	65	8	1	140	0	1483
s	234	128	3	9	277	823	192	0	13	0	27	0	2228
t	756	2	0	295	257	131	120	3	54	0	125	3	2343
u	4	81	0	306	256	263	6	3	0	2	3	1	255
v	46	0	0	0	2	0	1	1	0	0	5	0	0
w	159	0	0	13	45	2	0	0	0	0	3	0	326
x	1	47	0	0	0	23	0	0	0	5	0	0	21
y	64	9	0	9	44	5	4	0	3	0	2	1	1171
z	4	0	0	0	0	0	1	0	0	0	0	2	2
SPACE	721	588	42	494	1596	3912	134	116	1787	0	436	2	0

Figure A–4. **Bi-gram frequency table B_3, our own table generated from a corpus 10 times the size of that used for B_1 and B_2, consisting of a mixture of informal and informal English (e-mail and classic novels).**

1st/2nd	a	b	c	d	e	f	g	h	i	j	k	l	m	n
a	2	1218	1968	2368	11	382	894	31	2274	32	695	4401	1419	9584
b	636	58	11	7	3595	3	6	1	419	58	0	1219	61	3
c	2272	0	283	4	3022	10	2	2915	687	0	950	805	2	0
d	854	14	3	313	3561	39	178	5	1816	16	10	264	36	126
e	3647	211	1587	6084	2768	957	390	121	658	14	162	2843	1715	7279
f	1019	1	0	4	963	523	2	1	1380	0	0	280	1	5
g	571	0	6	6	1966	8	178	1716	653	0	0	336	35	176
h	7858	22	2	4	16596	9	0	0	5444	0	0	43	41	48
i	854	407	2903	2007	1314	814	1428	2	15	7	281	1930	1821	12853
j	43	0	0	8	154	0	0	0	0	0	0	0	0	0
k	50	1	0	3	1606	28	13	3	526	2	0	60	5	432
l	2464	30	26	1760	4571	395	151	4	2577	0	181	3711	53	9
m	2886	360	18	4	4331	12	1	1	1399	1	1	21	356	57
n	1183	18	1924	7392	3928	304	5759	23	1450	53	587	467	31	401
o	283	452	709	837	182	5093	422	67	714	15	581	1312	3360	8544
p	1373	6	4	16	2122	27	7	296	446	0	7	1144	32	0
q	0	0	0	0	0	0	0	0	22	0	0	0	0	0
r	2683	87	267	1373	9285	98	257	109	2860	12	505	295	716	721
s	1462	41	486	53	5033	111	19	1973	2564	14	273	279	342	94
t	2364	6	229	2	5364	25	1	18290	5137	0	3	810	62	48
u	571	570	884	249	611	89	858	1	547	0	15	2174	509	2087
v	534	0	6	0	5429	0	0	0	966	0	0	2	2	0
w	2620	4	5	17	2325	4	0	2772	2475	0	6	76	3	602
x	203	0	124	0	94	5	1	5	218	0	0	2	1	0
y	36	40	30	30	503	9	1	4	214	0	0	60	94	49
z	40	0	0	0	193	0	0	0	43	0	0	18	0	0
SPACE	15837	5825	5611	3963	3009	5348	2053	8139	8886	397	569	3364	6013	2993

1st/2nd	o	p	q	r	s	t	u	v	w	x	y	z	SPACE
a	24	1024	3	5042	5285	8404	559	1924	430	106	1715	52	4038
b	954	5	0	620	280	132	1082	59	1	0	652	0	85
c	3299	0	20	608	82	1745	742	0	0	0	132	11	278
d	1638	35	0	534	606	7	453	115	18	8	318	0	13169
e	276	774	225	10955	5514	2867	23	1360	526	960	616	29	25931
f	2731	2	0	1121	18	476	517	0	2	1	68	0	5246
g	815	0	0	712	216	48	340	0	0	0	56	1	3899
h	2824	0	0	443	50	1121	334	0	6	0	194	0	3634
i	2950	305	76	1392	6474	6267	31	1150	0	115	1	204	37
j	125	0	0	0	0	0	337	0	0	0	1	0	0
k	14	0	0	1	297	6	6	0	13	0	37	0	1283
l	1976	138	0	80	536	422	555	112	75	0	2487	1	3347
m	1626	1083	1	1	428	5	671	0	2	0	1305	0	2034
n	3137	41	32	39	1926	4524	294	205	29	15	566	6	9940
o	1927	927	2	6134	1359	2523	7904	704	2519	54	118	13	6488
p	1761	651	0	1669	286	316	380	0	8	0	64	0	903
q	0	0	0	0	0	0	846	0	4	0	0	0	4
r	3521	282	2	846	1636	1397	629	245	133	0	1570	0	7160
s	2119	712	59	13	1962	5112	1275	3	200	0	227	0	13916
t	6280	15	0	1855	1239	1408	1043	2	589	0	759	5	14050
u	55	1328	3	3135	2305	2832	3	5	17	18	27	15	1621
v	159	0	0	1	1	0	8	0	0	0	30	0	23
w	1471	3	0	218	181	4	4	0	5	0	9	0	1303
x	4	232	1	1	0	221	3	0	4	2	11	0	150
y	2456	73	0	4	450	115	8	0	27	0	0	4	7491
z	5	0	0	0	0	0	12	0	0	2	4	25	26
SPACE	9289	3657	396	2787	9383	20491	1969	1028	9162	25	2548	16	8079