A Study of Eye and Finger Behaviors for Text Input in Mobile Interfaces

by

Xinhui JIANG

Student ID Number: 1218002

A dissertation submitted to the
Engineering Course, Department of Engineering,
Graduate School of Engineering,
Kochi University of Technology,
Kochi, Japan

For the degree of
Doctor of Philosophy

Assessment Committee:

Supervisor: Prof. Xiangshi REN
Co-Supervisor: Associate Prof. Toru KURIHARA, Kochi University of Technology
Co-Supervisor: Associate Prof. Yukinobu HOSHINO, Kochi University of Technology
Associate Prof. Yoshiaki TAKATA, Kochi University of Technology
Prof. Shinichi YOSHIDA, Kochi University of Technology

March 2021

ABSTRACT

A Study of Eye and Finger Behaviors for Text Input in Mobile Interfaces

Text input on mobile devices is getting prevalent. Users of different levels of ability and expertise are learning and using soft keyboards. However, compared with typing on a physical keyboard, users face more challenges on mobile devices in eye and finger control. First, soft keyboards in mobile interfaces are generally smaller than physical keyboards and built without tactile feedback. Those features and constraints make users rely more on visual guidance while moving their fingers towards key buttons. Adding that proofreading on the text input area is also required, users have to shift their attention frequently across the text input area and the keyboard, which takes time and cognitive resources. Second, as the keys are mostly smaller than fingertips, users are likely to commit more errors due to occlusion problems. Third, users have to adapt their operations to mobile devices and press keys with only one or two fingers. While typing on physical keyboards can be guided by the 10-finger touch-typing technique, there is no standardized typing method on smartphones. To achieve a relatively faster typing speed on mobile devices, users have to develop their typing strategies. With the improvement of typing skills, users learn how to control their eye and finger movements to adapt to the task and constraints.

Existing text input designs in mobile interfaces are mostly based on the qwerty layout on physical keyboards and design guidelines from studies analyzing key-pressing logs and touchpoint distributions. However, it is still unclear how the touchpoints were made and how users managed to adapt to the mobile typing environment, specifically 1) how attention resource was allocated between tasks like key-searching and proofreading, 2) how fingers were guided with different level of typing proficiency to press the keys, 3) how errors were made and corrected, and 4) how the eye and finger behavior changed from being a novice to a more skilled typist. Answering those questions can help us re-think the current typing strategies and keyboard designs, and finally, provide more in-depth guidelines for learning and designing the interface for text input. To understand the typing behaviors on mobile devices, we conducted four studies, starting from basic pointing tasks to the learning process and expert typing

performance, to gradually unveil the underlying mechanism of attention and movement control in typing.

As typing is composed of sequences of pointing operations, we first conducted a study on a pointing task on the smartphone and summarized eye and finger movement patterns. The results confirmed that finger-pointing movement on mobile devices was guided by gaze in four patterns: 1) finger following, 2) finger guiding, 3) locate and leave, and 4) finger guidance with peripheral vision. These behaviors vary across individuals. The findings of this study provided a basic understanding of the eye-hand coordination mechanism on mobile devices.

In the second study, we looked into the typing behavior through a transcription task on the smartphone. We captured finger and gaze movement in both two-thumb typing and index-finger typing conditions. We analyzed the typing behavior using both the basic metrics (e.g., typing speed and error rate) and detailed measurements based on gaze and finger tracking data. Results showed that adaptive attention sharing strategies play a significant role in typing. Even though mobile devices are relatively small, users could not monitor the keyboard and the text display simultaneously. A strategy must be decided to determine which part to give attention to and when. As for the speed-accuracy tradeoff, users had to decide on a strategy that strikes a compromise between the cost of not correcting errors early and the time lost in glancing at the text display when the fingers cannot be guided. The findings confirmed the importance of understanding eye and finger movement in typing tasks. We summarized the characteristics of relatively faster typists and compared them with those who are slow in typing. This study not only indicated how a faster typing speed could be achieved, but also supported further modeling and perdition of gaze and finger movement under different constraints of keyboard design by revealing the inner mechanism of attention sharing and movement control.

Understanding how user behavior differs from novice to skilled helps us find out the underlying learning mechanism and provide insights into how the typing performance could be improved. As it was hard to find users who are new to typing, in the third study, we controlled the typing proficiency by setting different levels of keyboard layout randomization. We captured the typing behavior under two conditions: statically randomized keyboard and dynamically randomized keyboard, and compared those behaviors with the data captured on the Qwerty keyboard in the second study. The Qwerty refers to a standard Qwerty layout on physical keyboards and soft keyboards on mobile devices. The statically randomized keyboard refers to a keyboard with randomized letter keys, which were kept unchanged during the experiment. The dynamically randomized keyboard refers to a keyboard with randomized letter keys, and re-randomized once the participant clicks on the keyboard. By controlling the keyboard randomization, we defined novice typists as someone who has no knowledge of where the keys are, intermediate typists as someone who has had some exposure to the layout, and skilled typist as someone who has had enough typing experience with the keyboard layout. Findings of the study not only revealed the differences in behaviors between novice and experienced typists, but also showed the adaption of eye and finger movement throughout the learning process.

Despite the challenges we discussed typing on a mobile device, some fast typists could still reach an average typing speed of 80 words per minute (WPM), which was twice as fast as the average. In the fourth study, we invited fast typists selected from a web-based typing test and studied their typing strategies. In order to capture their typing behavior in an everyday setup, we developed a web-based experiment application and asked participants to transcribe target sentences on their own mobile devices. The findings of this study showed that fast typists focused on the text input area most of the time, which challenged the original understanding of how people type on mobile devices in study 2. Compared with touch typists on a physical keyboard who were trained to type with ten fingers and focus on the monitor, the fast typists on mobile devices could also adapt their behavior to the constraints of the soft keyboard and apply a similar strategy. Thus, this raises a discussion of how typing strategies could be shared among different devices and environments.

In summary, this dissertation looks into the typing behavior from the perspective of eye and finger movement, which revealed the overlooked details of typing on mobile devices. The contribution can be summarized as follows: 1) in-depth findings on eye and finger movement control during typing on mobile devices, which to our knowledge, has not been done before, 2) theoretical foundation of human skill learning by systematical investigation of normal typists, fast typists and the transformation process from novice to a skilled typist, 3) detailed dataset supporting the development of models for eye and finger movement which can further contribute to the ability-based keyboard optimization, and 4) guidelines not only for text input design, but also for learning and improving typing skills. The conclusions drawn in this dissertation provide insights not only on how to leverage human capabilities for better typing performance, but also on how to provide text input designs for supporting it.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Background

Text input has become a regular activity on mobile devices. In the past decades, the implementation of mobile keyboards has transferred from using physical buttons on feature phones to using the integrated touchscreen keyboards with flexible forms on smartphones. Due to the convenience of the touchscreen and soft keyboard designs, users of different backgrounds and different levels of abilities started typing on mobile devices for purposes like instant messaging, emailing, or even inputting account information in applications. However, when typing first came into use, both the operations and the keyboard layout were designed for typewriters (i.e., physical keyboards). Compared with typing on a physical keyboard, users on mobile devices face more challenges, such as the occlusion problem and a lack of standard guidelines for high-performance input.

Research conducted on improving typing performance considered changing the keyboard layout [1] and adding algorithms to correct the deviations of the touchpoints [2]. Those solutions were proposed based on findings of typing studies that captured touchpoints and key-pressing logs on mobile devices [3–5]. Although the outcome of the studies could be helpful, it is still unclear how the touchpoints and the key-pressing behaviors happened and how users made decisions during typing (e.g., when to proofread). Understanding the process of attention shifts and finger movement control during typing could be seen as a new perspective of looking at the typing mechanisms which contribute to human behavior modeling on mobile devices. To capture the complex and changing movements during typing, we build a tracking system with an eye tracker, a motion-tracking system, and a smartphone. Then we looked into typing behavior from perspectives, including simple pointing tasks, typing tasks with different postures, the learning of typing, and features of the fast typists. Systemic analyses were conducted based on attention and finger movement control. We not only report the relationship between eye and finger behaviors and the typing performance (e.g., typing speed and error rate) but also summarized guidelines for both the users who want to improve their typing skills and the practitioners who aim to improve the design of typing techniques and interfaces.

1.2 Motivation and Objectives

Although previous studies have discussed typing performance based on typing logs and touchpoints, there still remain questions that need to be answered based on attention and finger movement control: 1) what strategies are there for typing on mobile devices, 2) how do users learn a new keyboard layout, and 3) how do typists achieve fast typing speeds. As users' abilities differ individually, analyzing the detailed movements helps to identify the reasons for the performance differences. In this dissertation, we describe the tracking systems built for typing experiments, present studies aiming for capturing human typing behaviors in different situations, and summarize the strategies users applied in each case. We aim to understand the details of attention and motor control in typing tasks, and their connection with the typing performance metrics like speed and error rate. The findings offer insights for understanding the underlying coordination of visual and motor systems during typing. Detailed movement data captured in the studies could contribute to the modeling of eye and finger movement on mobile devices, which is regarded as an essential foundation in the development of an adaptive user interface with holistic consideration of the users' decisions and behaviors.

1.3 Dissertation Overview

This dissertation is concerned with studies of attention and finger movement control on mobile devices. Chapter 2 presents a literature review on typing performance, keyboard designs, and supporting algorithms, followed by attention and finger control in typing. Chapter 3 presents a study on eye and finger behavior during pointing tasks. Chapter 4 looks into the typing behavior through a transcription task on the smartphone, and presents results and findings in both two-thumb typing and index-finger typing conditions. Chapter 5 presents a study that controlled users' typing proficiency by setting different keyboard layout randomization levels and explained the change of eye and finger movements in the learning process. Chapter 6 presents a study on fast typists who reached twice the typing speed of regular users. Their strategies during typing are listed. Finally, chapter 7 summarizes the work and provides general conclusions and future research directions.

CHAPTER 2

LITERATURE REVIEW

2.1 Typing Performance

2.1.1 Typing Speed

Typing on a touchscreen keyboard on a mobile device differs from typing on a physical keyboard, as the keys are generally smaller and unable to provide tactile feedback. Usually, users have to hold the device by hand, leaving only one or two fingers for key-pressing. In order to improve their typing performance, users have to adapt their typing behavior to those constraints and develop their own typing strategies. In the perspective of posture and finger usage, studies on mobile typing reported a higher typing speed with two fingers than with only one finger [3, 6].

Apart from posture, there are other factors that relate to typing speed. A study [7] on the effects of the key sizes reported a growing typing speed from 23 to 27 words per minute (WPM) with increasing virtual keys sizes from 13 to 22mm. Similarly, compared with using a watch-sized keyboard, users could reach a higher typing speed by using a phone-sized keyboard [8]; compared with portrait mode (25 WPM), typing in the landscape mode with two thumbs tended to be faster (29 WPM).

Typing speed can also be affected by internal language processing. A text composition task takes more time than a transcription task [9]; memorized typing is faster than copy typing [10]. Keeping a high typing accuracy takes time. A recent study [11] showed that typing speed dropped from 60.16 to 40.91 WPM when the participants were required to be more accurate. The aversion of errors also reduced the upper bound of expert typing speed on a mobile soft keyboard from 33.26 (SD = 3.58) to 30.88 (SD = 3.58) WPM, a 7.21% difference [12]. However, it was reported that a higher autocorrect accuracy could lift this upper bound. Expert typists adjusted their speed-accuracy tradeoff to type faster and make more errors as the accuracy of autocorrect increases [13].

2.1.2 Error Rate

Generally, typing on a touchscreen keyboard was less accurate than on a physical keyboard, no matter which touchscreen device was used [10, 14]. Although typing with two thumbs on mobile devices was faster than using one finger, it also led to more errors [3, 15]. It was also reported that the keys with the highest error rate were more distant from the dominant hand [15].

In order to type more accurately, a larger keyboard could be used. A study [8] showed that compared with using a watch-sized keyboard, a phone-sized keyboard could reduce the error rate. But further larger keys might not continuously contribute to this trend; the error rate increased from 7.7% to 11.1% as the key size grows from 13mm to 22mm. Other measures like showing the touchpoint location [5] and reducing mobility [15] can also reduce the error rate in typing. However, as visualizing touchpoints with dots reduced the error rate (between 3.8% and 18.3%), it still decreased the typing speed up to 5.2% [5].

Apart from interface design factors, the task requirements also relate to the error rate. A recent study [11] showed that the error rate dropped from 12.2% to 0.8% as the requirement changed from being fast to being accurate. Compared with memorized typing, the users are more likely to correct the errors in copy typing, leaving a higher corrected and a lower uncorrected error rate [10]. A more challenging phrase for typing also leads to a higher error rate [8]. As for attention, a higher error rate was observed when participants were asked to use their peripheral vision for typing on a split keyboard on the tablet. As for typing on a physical keyboard, it was reported that faster typists make fewer mistakes [16]. In particular, they make fewer substitution errors, whereas insertion errors correlate less with performance.

2.2 Keyboard Design and Optimization

The most widely used keyboard layout is Qwerty, which was designed by Sholes in 1873 to minimize jamming in mechanical typewriters by placing common digraphs (i.e., consecutive letter pairs) on opposite sides of the keyboard [17]. Although the problem no longer exists on the devices we are currently using, the Qwerty layout was kept and applied on almost all kinds of text input devices and interfaces. As input techniques like gesture input and auto error correction emerge, the problems of the Qwerty layout appeared. First, it led to ambiguity in the keyboard decoder [18, 19]. In gesture input, users are asked to make a stroke through all the keys of the words. Such gestures are then analyzed and mapped to identical key-pressings. However, as vowel buttons such as u, i and o, are physically close to each other, many pairs of words share identical or similar gestures (e.g., pit vs. pot). Smith, Bi, and Zhai [19] analyzed the gestures of a 40000-word lexicon and found that 6.4% of words have another word with an identical gesture on the Qwerty layout. Such ambiguity problems bring trouble for decoding. Second, it requires a long travel distance in gesture typing and one-finger touch typing. As the Qwerty layout was designed by physically separating the digraphs from each other, the finger has to travel across the whole keyboard frequently when using one finger for input.

To address the problems mentioned above and improve the typing experience, researchers have been searching for new keyboard layouts, considering the constraints of finger travel distance, tapping ambiguity, and bigger key sizes with letter grouping [19–23]. However, although the optimized layouts were proved to be better than Qwerty, they are still not widely applied [18]. The biggest challenge users face while using a newly optimized keyboard layout is learning. As the users are most likely to be familiar with the Qwerty layout, shifting to another keyboard layout decreased their performance abruptly, which was observed as a steep learning curve [21]. Users have to go back to the novice state and learn the new layout with practicing efforts to improve their performance. Such learning cost prevented most of the users from adopting optimized keyboards. As typing is an activity requiring continuous visual attention, understanding eye-hand behaviors on a new keyboard can help identify the obstacles users face and provide a better design.

2.3 Attention during Typing

Typing is a complex task which consists of coordination between visual and motor control. Specifically, typists have to frequently switch their attention to tasks like key searching, finger guiding, and proofreading during typing. The touch typing method was designed for improving typing performance by allocating specific groups of keys to each finger. Typists who were trained to control the keyboard with ten fingers tend to keep their attention mainly on the monitor while typing on a physical keyboard [24]. Similarly, a study applying self-report showed that ten-finger touch typists pay more attention to the template in copy typing, more attention to the screen in free typing, and less to movement-related information (fingers and keyboard) than idiosyncratic typists (non-touch typists) [25]. In the perspective of changes in the performance with different attention allocation styles, a study [26] compared monitor gazers and keyboard gazers, and found that monitor gazers typed significantly faster and used cursor keys less. However, a more recent study [24] proved that non-touch typists could also reach the same typing speed as touch typists who were trained to keep their attention on the monitor during typing.

As for touchscreen devices, visual guidance was also found during internet activities on a tablet. It was reported that the gaze reached the target earlier than the finger, starting 309ms before the finger moment [27]. This control under uncertainty of the target location was confirmed by [28] that the handheld target locations for saccades could be updated by using proprioceptive information from the limb. However, this proprioceptive input is not as accurate as visual input in guiding saccades. Although there exist studies and discussions on finger control or attention during typing on the physical keyboard, it is still unknown what strategies users apply while typing on a much smaller device with a touchscreen.

2.4 Eye-hand Coordination

Unlike mechanical movements under computational control, a human cannot accurately follow the pre-planned pathways of movements for aiming. Although proprioceptive information from the limb can be used to identify the rough position of the finger and hand, it is not as accurate as visual input in guiding the movements [28]. In touch interaction, it was found that the fixation arrives close to the target before the tapping moment, and stayed there to guide the finger [27]. While in a sequential aiming task, planning of the movement path was done with saccades before the initial finger movement [29].

As for cursor control in a computer setup, because the location of the cursor cannot be sensed physically by holding the mouse, visual guidance becomes even more important. Studies provide evidence that gaze leads the cursor to the targets most of the time [30–32]. In a web-based task, the gaze-cursor distance increases when the cursor is inactive, and decreases when the cursor is being actively used to examine, read, or perform an action [30]. In everyday work, the gaze-cursor distances are lowest from 100ms to 250ms before the click, then gaze (the last saccade) leaves the target before a click 7.7% of the time [31]. This means that the cursor movement can be continuously controlled a while after the gaze moves away. In a search and selection task, when the target location was unknown, the eyes lead the mouse by 300ms on average; when the approximate location of the target was known, the cursor often led the gaze in acquiring the target [33].

Typing on a touchscreen mobile device requires more visual guidance as the buttons are relatively smaller than in other scenarios. Users have to pay attention to the keyboards for key-searching and finger guiding against fat finger and occlusion problems. As the details of the eye and finger movement on the mobile device are not yet thoroughly investigated, to grasp a better understanding of human behavior in mobile touchscreen interfaces, a series of studies were designed and conducted, from the basic pointing task to typing under different constraints and user capabilities.

CHAPTER 3

STUDY 1: EYE-HAND COORDINATION IN POINTING TASK

Most of the interaction scenarios on mobile touchscreen devices involve finger-touch operations (i.e., pointing). Such operations gained popularity as they can be intuitively learned and conducted. However, since the flat screen cannot provide sufficient tactile feedback for finger navigation, users need to rely on visual information for guiding the finger to the targets. Although there exists plenty of work focusing on understanding and describing the properties of finger movements such as movement time and error rate, there lack information on how the finger was controlled on a mobile device, based on detailed movement records. Understanding eye-hand coordination during the pointing process will contribute to a better understanding of the user behavior, and furthermore, promote the modeling of the user behavior on mobile devices. We report the detailed data and findings for unveiling the user behavior patterns of visual-motor control during Fitts' task on mobile touchscreen devices.

3.1 Introduction

As the most basic task on mobile touchscreen devices, pointing operation is widely applied in interaction designs. The difficulty of the pointing task is well described by Fitts' Law as the logarithm of the amplitude (distance to the target) of the movement divided by the target width (the required accuracy) [34]. For multiple reasons, including occlusion and the fat finger problem, pointing operations on mobile devices is sometimes error-prone. Existing studies mostly focused on the touch properties, such as the spread of the touchpoints [3, 35] and the function area of the fingers on the screen [36]. Based on such works and accumulated knowledge, optimal target size was proposed [37], algorithms assisting touch operations (i.e., touchpoint correction) were developed and applied [38]. As for the existing work related to eye-hand coordination during the interaction process, eye and pointer coordination was studied [33]. They found that when the approximate location of the target is known, participants perform pointer movements that are initiated without eye guidance, possibly in parallel to an eye movement to the target item. In tasks requiring a visual search for a target item prior to selection, users parallelize search and pointer movements. However, those findings cannot be easily applied to touch operations as the user was monitoring the mouse movement based on the visual feedback on the screen, which minimized the occlusion problem. The processes of finger movements and attention shifts during the touch operation were seldom mentioned or studied. Our goal is to reveal how humans control their eyes and hands to complete the pointing task on mobile touchscreen devices and how their strategies change according to the constraints such as target size, movement direction, movement distance, etc.

Although eye-hand coordination was investigated in different scenarios, such as using a mouse on a PC [30, 31], drawing [39], and even during driving [40], limited attention was focused on mobile devices. The possible reason is that building a tracking system based on eye movement and finger tracking with synchronized time is relatively hard. The most related work mentioned was conducted on a tablet [27]. However, it was claimed that the pointing performance could be affected by the size of the touchscreen [41]. Thus the results cannot be simply shared among different devices. In this study, we built the tracking system with a Dikablis eye tracker, a Vicon motion tracking system, and a smartphone. By pre-defined button-pressing actions, we synchronized the time among three devices and captured real-time tracking data. Our work aims to enhance the understanding of user behavior during a Fitts' pointing task by generalizing the patterns of the gaze and finger movement behavior. Findings will provide design references for the mobile touchscreen devices.

3.2 Method

3.2.1 Participants

Twelve volunteers were invited to participate in the experiment (4 males, 8 females). The age of the participants varied from 20 to 42, with an average of 23.8 years (SD = 6.1). All participants reported to have been using the smartphone on a daily basis; five of them used a smartphone more than 4 hours a day. The screen size of their smartphone varied from 4.7 to 6.1 inches, with an average of 5.3 inches (SD = 0.6). We asked about how participants usually hold their smartphone by providing four options (multi-selection): a) one-handed operation with the thumb of the dominant hand, b) operating with the thumb of the dominant hand and hold the device with the non-dominant hand, c) operating with two thumbs, d) operating with the index finger of the dominant hand and hold the device with the non-dominant hand. The numbers of participants reporting to match the situation of a, b, c, and d were seven, five, five, and seven.

3.2.2 Experiment Design

The experiment was designed for understanding the eye-hand coordinative movement during pointing tasks. The purposes are to 1) bring the pointing operations out of the complex contexts (typing or menu operation) and understand the mechanism of eye-hand movement in a simple task; 2) confirm the effects of target properties (i.e., target size, target distance, and moving direction) on the pointing performance and eye-hand movements.

Figure 3.1 Start positions (crosshair) in the pointing task (top, middle, bottom).

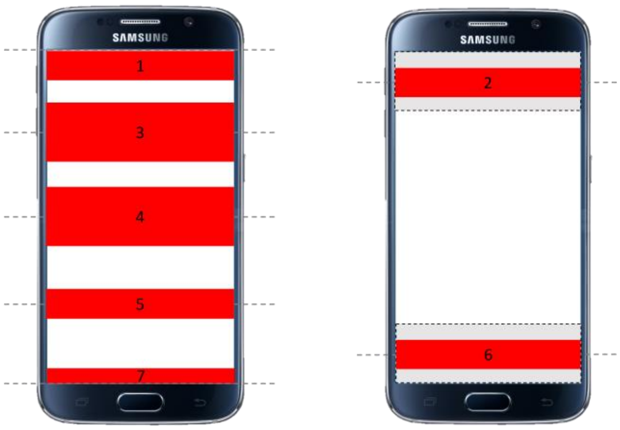


Figure 3.2 Target positions in the pointing task (1: top, 3: ¼ screen, 4: ½ screen, 5: ¾ screen, 7: bottom, 2: unified top, 6: unified bottom).

The pointing task was designed as pairwise operations: pointing from the start point to the target is defined as one trial. We want to unify the gaze points' location while participants were pointing at the starting positions and designed the start positions as crosshairs. We asked the participants to try to hit on the center of the cross in order to ensure that their eyes are looking at it. The starting positions are designed as top, middle, and bottom (Figure 3.1). We used three target sizes, which were 454 pixels (the edge of the target is visible while the finger is touching it), 227 pixels (icon size), 114 pixels (small). For identifying the target positions, we vertically divided the screen into four parts (Figure 3.2, left). Targets 3, 4, and 5 were located on the division lines. Targets 1 and 7 were at the top/bottom borders of the screen. However, as the centers of the border targets were not unified, we define target positions 2 and 6 by putting the biggest target (454 pixels) on the top/bottom borders and use the centers of that target as target positions (Figure 3.2, right). We are also interested to see if the results show a significant difference between those two kinds of border targets.

When the pointing task starts from the top or bottom of the screen, five target positions could be tested. When starting from the middle of the screen, the number of target positions became six. We name the combination of start points and target positions as conditions. Thus, for the pointing experiment, each participant will conduct: (5+5+6) conditions $\times$ 3 target sizes $\times$ 5 repeats = 240 trials.

3.2.3 Setup

We conducted an experiment by using a Vicon motion tracking system (12 cameras) for finger and smartphone tracking, a Dikablis eye-tracking system for capturing the gaze data, and a Samsung smartphone for capturing touch operations (Figure 3.3). Markers for tracking the movement of the smartphone were stuck to a marker holder, which was pasted on the smartphone. One marker was pasted on the participant's index finger of the dominant hand. Other markers were pasted on a glove, which was worn on the participant's dominant hand. For a better understanding of the coordination of eye and finger movement, we developed an embedded mechanism for synchronization across the devices used.

Figure 3.3 Experimental setup.

3.2.4 *Synchronization among Devices*

We developed the pointing application with a synchronization process before the experiment. During synchronization, participants were shown the interface of the application with four buttons (Figure 3.4). They were instructed to click on the four buttons following the numbers on them in ascending order. While the buttons were clicked, a blue rectangle flashed on the screen (Figure 3.5). At this moment, the eye tracker, motion tracker, and smartphone captured the button-pressing behavior in different ways: the eye tracker captured the blue flash on the screen with the front-faced camera; the motion tracker logged the finger movement trace with the 3D coordinates of the finger and smartphone; the smartphone captured the button-pressing operation by “touch down” events in the application.

During data processing after the experiment, videos from the eye tracker were processed for recognizing the blue flashes, which indicate the moment of button-pressing; finger coordinates were analyzed for identifying the lowest finger location from the screen during the button-pressing. The timestamps of the two events were then matched with the timestamp of button-pressing operations on the smartphone for synchronization.

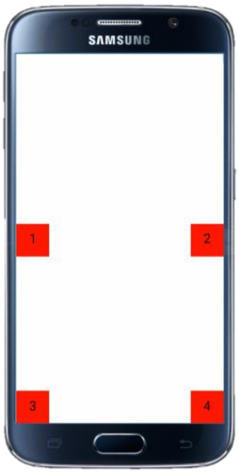


Figure 3.4 Application interface for synchronization.

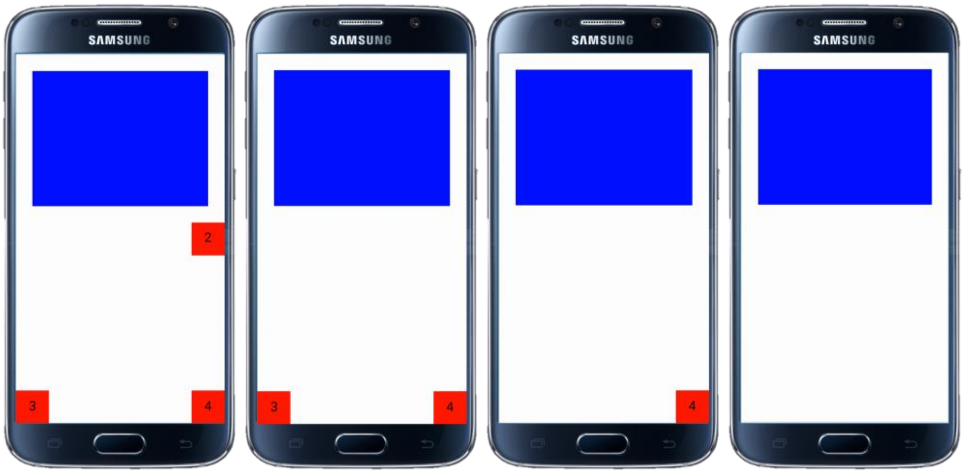


Figure 3.5 Application interface while clicking on the buttons for synchronization.

3.2.5 Procedure

First, the subjects were asked to sit in a chair beside a table within the range of motion tracking (Figure 3.3). We explained the purpose, procedure, and task of the experiment and asked them to fill in the consent form. We put on a glove with serval markers on the participant's dominant hand and asked them to hold the smartphone with the other hand. Then we put the eye tracker on the participant's head, and adjust the position. After confirming that all the device works properly, we started the practice session for 5 minutes. Before the experiment session, we performed calibration with the Dikablis eye tracker software (D-Lab). We then started the synchronization process by asking the participant to point on the four buttons on the screen in ascending order. After synchronization, the experiment interface was shown.

As for the experimental task, we asked the participant to point at the targets shown on the screen one at a time. There were two types of targets, one of them is a crosshair-shaped target, indicating a start point: when participants see the crosshair-shaped target, they were supposed to look at it and click on its center. The other type of target is a rectangle-shaped red target, indicating the target for pointing. This pointing task refers to a typical Fitts' task on the smartphone. Participants were asked to click on the red target as fast as accurately as possible. Each trial refers to pointing from a crosshair-shaped target to a rectangle-shaped target. The experiment contains 240 trials of pointing operations, which lasts for around 5 minutes.

3.3 Results

To understand the performance and behavior during the pointing task, we conducted data analysis on the following aspects: pointing performance, the patterns of eye-hand coordination during the pointing task, and checked the relationship between eye-hand coordination and pointing performance.

3.3.1 Pointing Performance

As for controlling factors, the location and size of the pointing targets could also decide the direction and distance of the finger movement. According to Fitts' law, target size and distance affect the time spent for pointing operations. Thus, we firstly analyze the effects of the target size, target position, movement direction, and movement distance on pointing time and error rate. Before statistical analysis on the pointing performance, we filtered the data by dropping outliers with pointing time landed outside the mean ± 2 * std. We checked the effect of condition (the combination of start and target positions) and target size on time spend for pointing behavior using two-way repeated-measures ANOVA. Results showed that both condition ($F_{15, 165} = 15.72$, $P < 0.001$) and target size ($F_{2, 22} = 33.37$, $P < 0.001$) has significant effect on the time spent for pointing. There is a significant interaction effect of condition $\times$ target size ($F_{30, 330} = 2.11$, $P < 0.001$) on the error rate of pointing. We also found significant effect of target size on error rate ($F_{2, 22} = 22.14$, $P < 0.001$). But there is no significant effect of condition ($F_{15, 165} = 0.95$, $P = 0.51$) and interaction effect of condition × target size ($F_{30, 330} = 1.12$, $P = 0.31$) on error rate.

(1) Target size: time spent for pointing was measured as the duration from lifting a finger from the start point (“touch-up” event) to the contact of the fingertip to the screen for pointing on the target (“touch-down” event). As predicted by Fitts’ law [34], the time spent on pointing decreased as the target size increased. The error rate of the pointing task decreased as the target became larger. There is a significant difference between the error rate of the target of 114 pixels and others (228 pixels and 454 pixels), indicating that the optimal target size is bigger than 114 pixels (Figure 3.6).

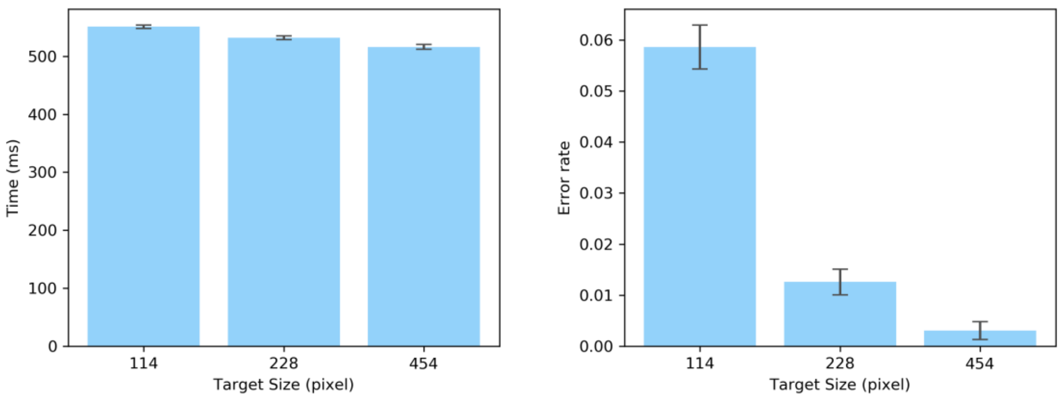


Figure 3.6 Pointing time (left) and error rate (right) for different target sizes.

(2) Target position (positions see Figure 3.2): the targets at the middle of the screen (position 4) led to the shortest time for pointing. As for the error rate, targets at the top or bottom of the screen led to a lower level of error rate (Figure 3.7). This generally fits with the speed-accuracy tradeoff: higher movement speed accompanies a lower accuracy for pointing.

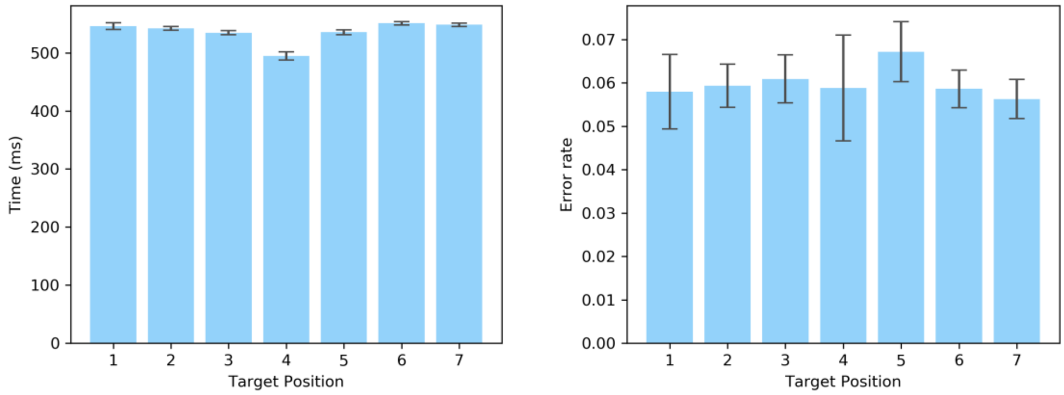


Figure 3.7　Pointing time (left) and error rate (right) for different target positions.

(3) Movement direction: during the experiment, participants moved their fingers upwards or downwards according to the locations of the starting point and the target. Results showed that, although the time for moving upward and downward was relatively the same, the error rate differed (Figure 3.8). Upward finger movements led to more errors than downward movements.

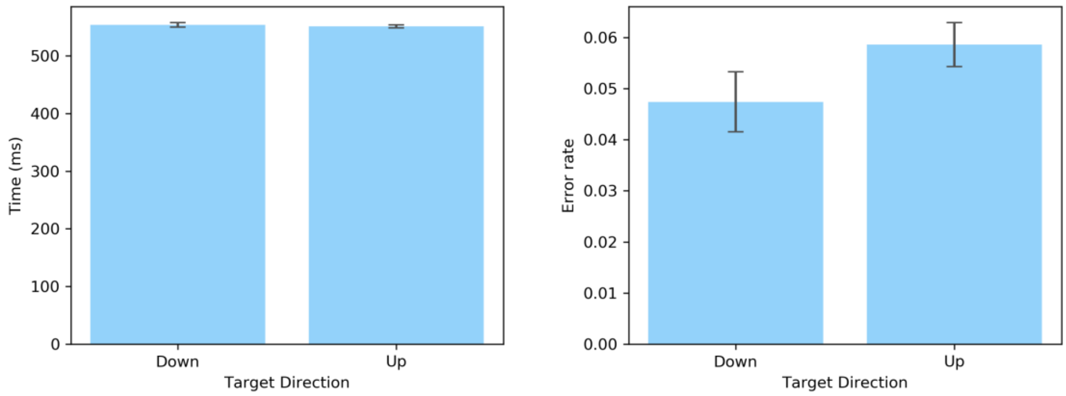


Figure 3.8　Pointing time (left) and error rate (right) for different target directions.

(4) Movement distance: the movement distance refers to the distance between the start point and the target. For a better understanding, it was transformed from pixels to centimeters. As shown in Figure 3.9, there is no specific trend of increase or decrease across different levels of movement distance. However, the time and error rate turned out to be the lowest while moving 5.1 cm for pointing.

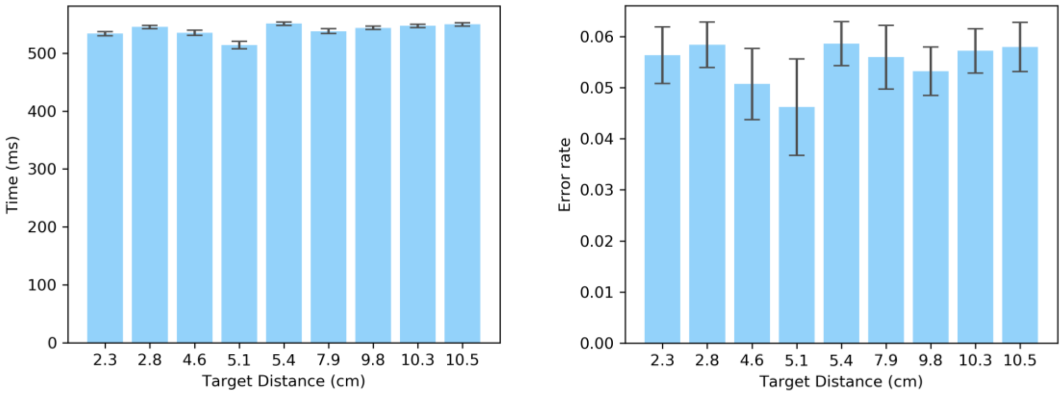


Figure 3.9 Pointing time (left) and error rate (right) for different target distances.

(5) Fitts law fit: we tested whether the pointing performance on a mobile touchscreen device follows Fitts' law [34]. Linear regression between movement time (MT) and Fitts's index of difficulty (ID) shows that MT = 418.99 + 51.01 $\times \log_2$ (D/W + 1) (Figure 3.10), in which D is the target distance and W is the target size. The R^2 = 0.8 for the empirical relationship between movement time (MT) and index of difficulty (ID = $\log_2$ (D/W + 1)). The relatively lower R^2 is probably due to smaller IDs on the size-limited screen.

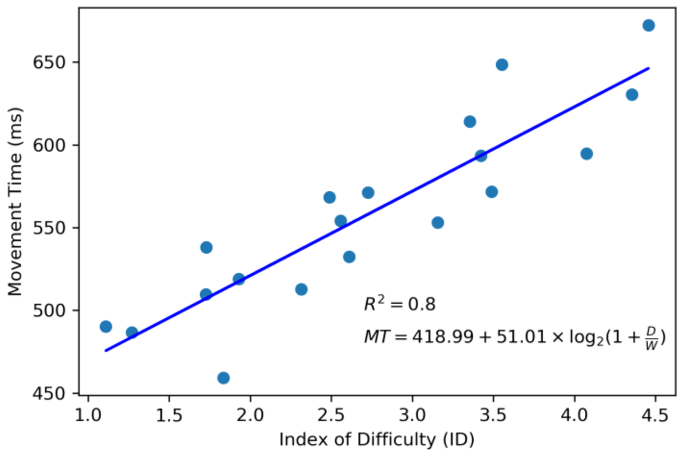


Figure 3.10 Linear regression with Fitts Law.

3.3.2 *Finger Movement Efficiency*

In order to understand how the finger was controlled during pointing movements, we defined finger movement efficiency as the target distance divided by the length of finger movement from the start point to the target in the 3D space, and visualized it under different target distances. Here, the target distance refers to the distance in centimeters from the start point to the target center. The result showed that as the target became further away from the start point, the finger moved more efficiently (Figure Figure 3.11 Finger movement efficiency for different target distance). The highest finger movement efficiency reached around 0.6, while the target distance is about 10cm.

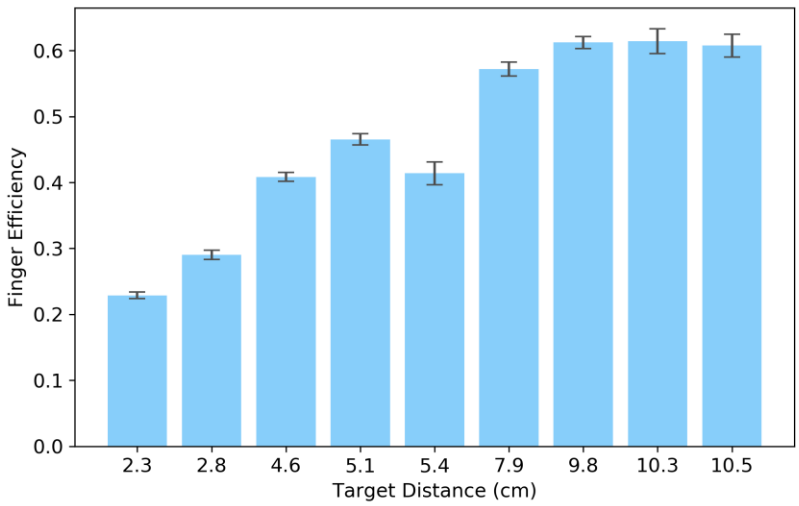


Figure 3.11 Finger movement efficiency for different target distances.

3.3.3 Eye-hand Coordination Patterns

To get a general understanding of how the eye and finer moved from the appearance of the targets to the finger touch on it, we drew line charts of the distance to the target from both eye and finger. From the perspective of eye movement, four patterns were found to appear frequently, which are described below:

(1) Finger following: after the appearance of the target, the gaze kept a similar distance from the target with the finger until the finger touched on the target. We assume that, in such a situation, the eye and finger followed each other while moving from the start point to the target (Figure 3.12).

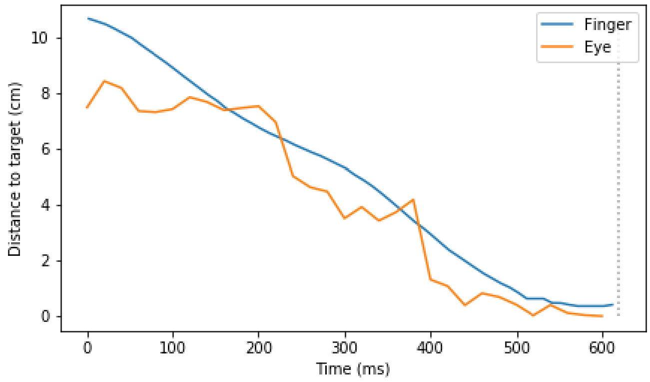


Figure 3.12 Example of “finger following”.

(2) Finger guiding: gaze first moved to the target and then waited for the finger to come. In such situations, although the gaze was not following the finger movements, it can still efficiently guide the finger by anchoring on the target, providing a reference for finger movement (Figure 3.13).

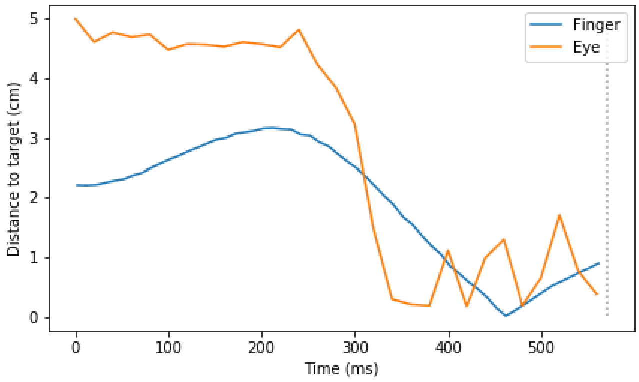


Figure 3.13 Example of “finger guiding”.

(3) Locate and leave: gaze first located the target, and left before the finger reaches the target. In such situations, the location of the target was well-confirmed before the finger came to touch. Thus the participant chose to move their gaze away in advance. It reflected that when the participant has this memory of the target location on the mobile device, they could guide their finger movements without focused visual attention (Figure 3.14).

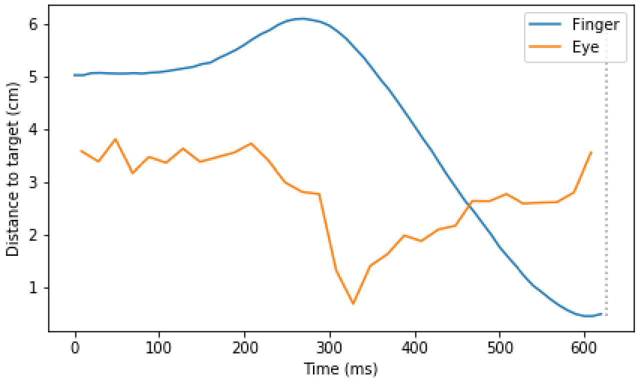


Figure 3.14 Example of “locate and leave”.

(4) Finger guidance with the peripheral vision: gaze stayed away from the target without coming close to it. Finger guidance with peripheral vision can be efficient if the target is big enough. In such situations, the participants trusted the accuracy of the pointing operations, and chose to move their fingers without focused attention on the finger or target (Figure 3.15).

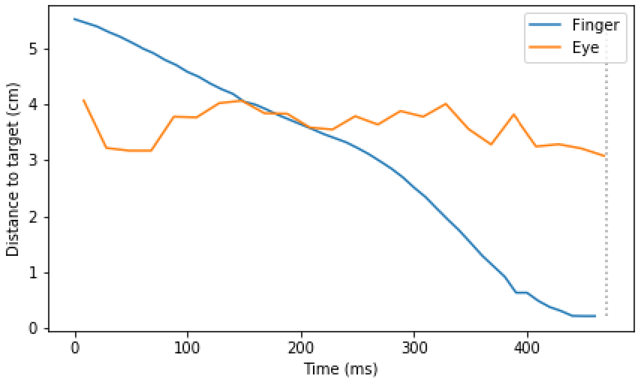


Figure 3.15 Example of “finger guidance with the peripheral vision”.

3.3.4 Eye-hand Relationship and Pointing Performance

For a better understanding of how eye-hand coordination affects performance, we drew scatter plots of the effect of eye-hand distance on the pointing time based on each pointing movement (Figure 3.16). There existed a trend of increasing pointing time while the eye-hand distance increased. This indicated that the coupling of eye and finger movement during pointing contributed to faster movements and task completion.

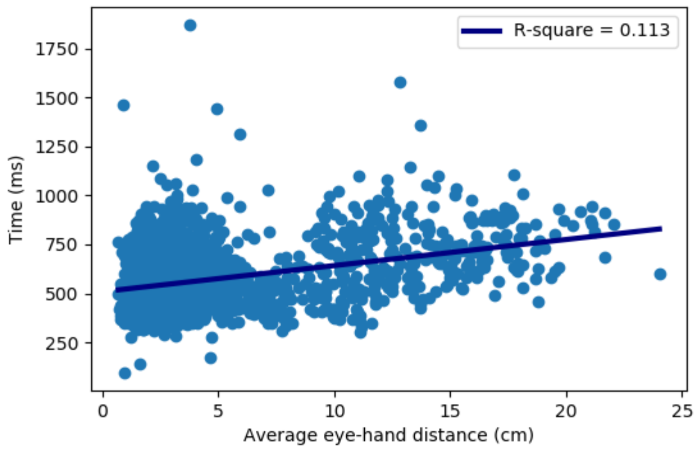


Figure 3.16 Scatter plot of pointing time for eye-hand distances.

3.4 Summary

Pointing operations are widely applied on mobile touchscreen devices in a variety of scenarios. Although there exists literature providing and developing models that described the pointing based on movement time and error rate, this study focused on the pointing behavior by analyzing the eye and finger movement. In this study, we built a tracking system with an eye tracker, a motion tracking system, and a smartphone, and captured the detailed data of eye and finger movement in each coordinate system. We transformed the coordinates into a uniformed coordinate system, as well as the same time system. We observed the user behavior from the perspective of pointing performance, eye-hand coordination pattern, and the effect of eye-hand coordination on the pointing performance. Findings showed that: 1) the optimal target size for pointing is bigger than 114 pixels, 2) targets at the middle of the screen are easier to be clicked fast and accurately, 3) moving upwards for pointing led to more error than moving downward, 4) there were four patterns for eye-hand coordination, which were: finger following, finger guiding, locate and leave, and finger guidance with peripheral vision, 5) a shorter average eye-hand distance led to faster pointing operations. Those findings can provide a reference for the UI designers of mobile devices towards a better user experience and operational efficiency.

CHAPTER 4

STUDY 2: EYE AND FINGER MOVEMENT STRATEGIES IN MOBILE TYPING

Relatively little is known about eye and finger movement in typing with mobile devices. Most prior studies of mobile typing rely on log data, while data on the finger and eye movements in typing come from studies with physical keyboards. This study presents new findings from a transcription task with mobile touchscreen devices. Movement strategies were found to emerge in response to the sharing of visual attention: attention is needed for guiding finger movements and detecting typing errors. In contrast to typing on physical keyboards, visual attention is kept mostly on the virtual keyboard, and glances at the text display are associated with performance. When typing with two fingers, although users make more errors, they manage to detect and correct them more quickly. This explains part of the known superiority of two-thumb typing over one-finger typing. We release an extensive dataset on everyday typing on smartphones.

4.1 Introduction

This study presents new data on how people type on touchscreen devices. Present-day understanding of typing is rooted mainly in studies with physical keyboards [24, 42–44], which differ in a few important respects from touchscreen keyboards on handheld devices. The most obvious are size, hand postures, the role of the thumbs, and the lack of physical key switches. Everyday mobile typing is carried out on the move too. Little research exists on the implications of these factors for how people move their gaze and fingers, which is surprising, given the prevalence of these devices and general awareness that visual and manual strategies affect performance across a plethora of skilled activities, such as typing on physical keyboards [24, 26], driving [45, 46], and gaming [47, 48].

Our goal is to understand movement strategies in mobile typing. Visuomotor strategies are learned over experience and coordinate manual and sensory actions like hand and eye movements [49]. They are associated with the performance [50]. For typing, in general, a strategy is needed for coordinating the allocation of visual attention between the text display and the keyboard, and to guide the timing and speed of finger movements [16, 24]. A typing strategy must also regulate the speed and accuracy of aiming, which should adapt, depending on target sizes and the permitted rate of errors [35, 51–53]. Prior research on mobile devices, in particular, suggests that the sharing of visual attention may have an important role. It must be devoted to guiding the fingers because of the lack of tactile landmarks [54, 55]. At the same time, it may be needed to check the correctness of the text and of predictions/corrections made by any intelligent text-entry method. Visual attention is also required in searching for rarely used characters on the keyboard [55, 56].

To identify and quantify movement strategies in mobile typing and their relationship to typing speed, one must have synchronized eye and finger movement data. Such data are needed for revealing movement strategies that are hard to report verbally. Yet, previous work on mobile typing has relied primarily on log data for touch events, which is not ideal for in-depth studies of visuomotor strategies. Hence, several questions are open. Firstly, is visual attention tightly coupled with finger movement? With physical keyboards, faster typists generally keep their attention on the monitor more than the keyboard [24, 26]. One study of mobile devices with physical keyboards suggests that there are attention shifts between text display and the keyboard [57]; however, this phenomenon has not been described in detail for touchscreen typing. Secondly, what exactly initiates error correction? Do users detect errors directly after the erroneous keypress or much later? Thirdly, do strategies differ among common typing styles, such as one-finger and two-thumb typing? Two-thumb typing is known to be faster, with the gain attributed to the rapid alternation between the lateral sides of the keyboard [20]. For answers, data on eye and finger movement are needed. High-quality datasets are also critical for efforts in predictive modeling and intelligent text entry methods.

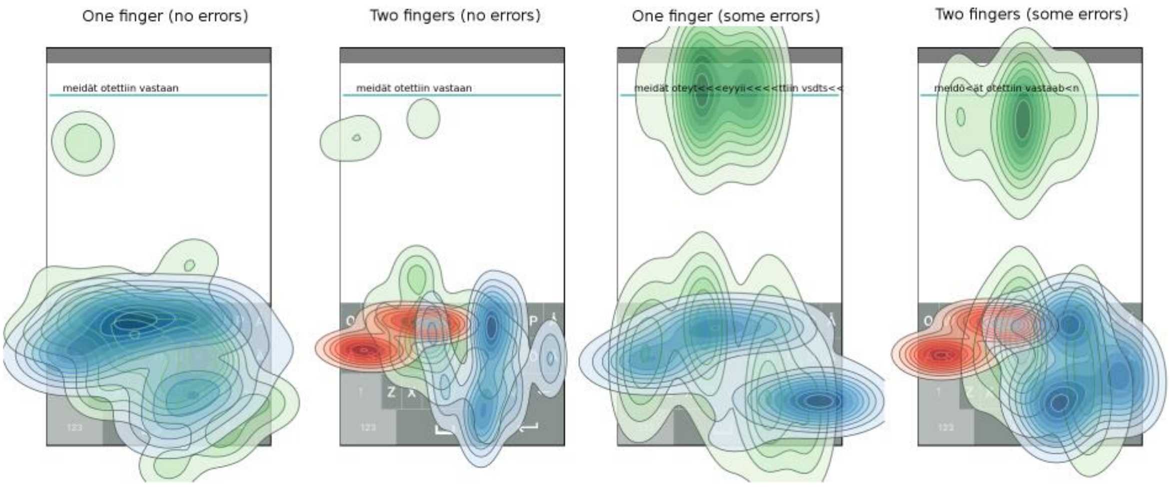


Figure 4.1 Illustration of our data as heatmaps of finger touchpoints (blue for right index finger or thumb, and red for left thumb) and eye movements (green). All present typing of the same sentence by a different participant: one (index) finger with no typing errors, two thumbs with no errors, one finger with errors, and two thumbs with errors. Glances at the text-entry area increase with the number of errors made, and error correction is visible as touches of Backspace. In two-thumb typing, visual guidance of the fingers is less in demand, so the gaze covers smaller areas of the keyboard.

Here, we report on findings from an exploratory study of transcription typing on a touchscreen device (N = 30), working with high-fidelity synchronized data from motion tracking, eye tracking, and on-device keypress logging. Our methodology for the study closely follows prior work on physical keyboard typing ("How We Type" [24]). To the best of our knowledge, the dataset presented here is the first of this type for mobile touchscreen devices. We report on eye movement, finger movement, eye-hand coordination, and predictors of typing performance with the use of both one and two fingers for touchscreen devices. Figure 4.1 illustrates different glancing behaviors along with finger touch in our dataset, for a given sentence typed with one and two fingers and with and without errors. We devote the rest of the paper to reviewing today's understanding of movement strategies in mobile typing, then reporting on our method and results. While we report many detailed analyses, our overarching finding is that gaze-deployment strategies are complex and much more important factors in typing performance than previously thought. We explain this and other findings in terms of how movement strategies adapt to the limited availability of visual attention. We also discuss the implications of the text entry system studied, which did not offer intelligent text entry techniques. The dataset is made publicly available.

4.2 Related Work

Typing is a complex visuomotor process that engages multiple cognitive, perceptual, and motor abilities [24, 58, 59]. This behavior has been a topic of research for almost a century, and typing on touchscreen devices for three decades [60–62]. Here, we discuss studies of physical and touchscreen typing, along with the contrasts they manifest. Papers on typing with a physical keyboard reported an average typing performance of around 50 words per minute (WPM) [16, 24]. Generally, typing speeds are lower for mobile devices, with reported averages between 36 and 41 WPM [3, 63].

4.2.1 Typing with a Physical Keyboard

Typing is a process carried out in phases, such as an input phase (grouping the to-be-typed text into chunks), parsing phase (decomposing the chunks into discrete characters), translation phase (converting characters into movement specifications), and execution phase (conducting the movements) [59, 64]. These phases are often interleaved, with the parallelism depending on a control hierarchy that is responsible for translating words into letters and motor plans [58]. Motor control strategies and ability affect typing performance. Expert typists can type quickly on account of the automatic translation of letters into motor plans, which can be executed quickly [65]. This is associated with consistent finger-to-key mappings, which, along with the preparation of the fingers, predict performance [24]. In addition, "rollover", wherein the next keypress is initiated while the previous key is still depressed, is prevalent among skilled typists especially; this improves typing performance [16]. Finally, alternating hands in the typing of bigrams is generally superior to typing them with a single hand [24, 59].

Typing also requires visual attention. Pointing movements often consist of a rapid ballistic and a slower corrective movement. Generally, the eyes and the pointing hand demonstrate a "pointing synergy", both moving towards the target at the same time, with the eye arriving earlier due to large saccade speeds [66]. However, the tactile feedback provided by a physical keyboard permits attending to the text-entry area for the majority of the time, resulting in fast detection and correction of typing errors [24, 26]. Typists who have studied touch typing and therefore have stable finger-to-key mappings do not have to glance down at the keyboard to search for keys [24, 44].

4.2.2 Typing on Mobile Touchscreen Keyboards

Touchscreen keyboards are generally much smaller than physical ones. Hence, most mobile typists use one or two fingers (generally thumbs) rather than the 3–9 fingers often used with physical keyboards [24, 63]. In a pattern similar to physical keyboards', the use of two fingers (generally thumbs) yields faster typing, via finger alternation and preparatory movements of the free finger [20, 67, 68]. To overcome the slower overall typing on touchscreen keyboards, several intelligent text-entry methods have been suggested [63, 69]. These include auto-correction of mistyped words, prediction of the next word, dynamic resizing of keys, touchpoint correction, and accounting for hand posture. In a recent logging study of mobile typing, more than 80% of participants used some sort of intelligent text-entry aid [63]. In the interest of starting from a simpler visuomotor-cognitive problem, we here focus on the non-aided case.

In addition to being smaller than physical keyboards, mobile touchscreen keyboards lack the tactile feedback of physical keys [70]. The fingers, lacking a physical reference point, need to be constantly monitored and guided by visual attention. Therefore, attention-sharing strategies differ between physical- and touchscreen-keyboard typing. Glances at the text-entry area permit proofreading of the text entered but hinder visual guidance of the fingers, reducing typing speed [63]. However, undetected errors are costly; when detected later, mistakes require more steps and time to correct. Users are known to slow their typing in response to errors [12], and the strategic finger speed-accuracy tradeoff and proofreading frequency have a large impact on text-entry performance [53, 63, 71].

Some studies have investigated the role of eye movements in non-typing touchscreen interactions. On internet-based tasks on tablets, gaze has been observed to precede touch with similar spatial and temporal features as observed with the mouse, but with individual differences [27, 72]. Interaction between gaze and touch has also been utilized to adapt UIs with the help of a predictive model [73]. On tablets with split keyboards, it is often enough to attend only the text entry area and use peripheral vision to guide fingers [74].

4.2.3 Theories and Models

Typing models make predictions such as transcription time, inter-key-interval, and the number of errors. Work thus far has focused mainly on modeling typing on physical keyboards, with fewer attempts relevant for mobile typing. Arguably the most popular statistical models are based on Fitts' law, which models aimed at movement performance [34]. After calibration of its empirical parameters to touchscreen pointing, it can predict performance over a range of layout conditions [35]. However, this family of models can only approximate the skilled typing behavior achievable after extensive practice [75]. Absent from these models are the adaptive strategies that govern the distribution of visual attention and describe the visual guidance of finger and proofreading activity or the frequency of proofreading [12, 53, 55, 56, 71, 76].

Keystroke-Level Models (KLMs) break task execution into operations, such as recall, pointing, homing, and attention shifts [57]. However, being sequential models, KLMs do not cover parallel movement, learning, or the role of attention. Nor do they predict how the interface or the user's abilities affect the choice of typing strategies. Finally, simulation models are step-by-step programs emulating the cognitive and physical steps involved. One recent model covers 12 operations or production rules: creating a mental representation of the task, visually attending the target, pointing at the target, confirming that the task is done, etc. [54]. The operations are simulated with a cognitive architecture, which computes their execution times and links together the separate cognitive modules, such as memory and attention. Predictions can be generated for a wide range of task conditions. The model predicts how typing performance is influenced by, for instance, changes in the number of keys or features such as their size.

The choice of movement strategy is very difficult to model with production rule-based cognitive architectures due to the sheer number of possible strategies. Recent research has turned to computational rationality [77, 78] to simulate strategic adaptation of gaze to the task environment [53, 56]. There, typing performance is modeled as an adaptation of eye and finger movement to the constraints of the human visuomotor system and the interface. At the moment, however, this class of models does not fully cover typing phenomena, including parallel finger and eye movement; one hindrance has been the absence of a rich dataset.

4.3 Method

We designed an experiment to obtain a rich dataset of movement strategies in a transcription task. Participants typed representative everyday messages and were instructed to correct typing errors. This is consistent with other research, where the instruction is often to type “quickly and accurately” [16, 63] or to correct errors upon noticing them [24]. We collected data for typing with the index finger and with two thumbs, the most common typing styles [63]. We used a Qwerty keyboard without intelligent typing aids to establish a dataset of baseline typing phenomenon. All data were synchronized in time, and all positions were registered in a single coordinate system.

4.3.1 Participants

We recruited 33 subjects. Because of gaze-data loss (device error), the number of participants decreased to N = 30 (18 females; age range 18–45, M = 25.5, SD = 5.9). Three participants were left-handed (1 female). All participants were native Finnish-speakers and had a normal or corrected vision (correction strength between -4 and +4). All reported using Finnish in their typing, with most using computers (desktop or laptop) several times a day (two reported using only a few times a month). Also, most used touchscreen devices (mobile phones or tablets) several times a day (one reported once-a-day use). The participants reported spending, on average, 16.7 hours (SD = 16.4) a week typing on a physical keyboard and 11.6 hours (SD = 8.0) on a mobile software keyboard. In our study, we observed typing performance of between 14.9 and 58.4 WPM for two-thumb typing and 19.1–33.3 WPM for one-finger typing. Each participant was compensated with two movie tickets (total worth about € 20) for their time.

4.3.2 Experiment Design

Each subject typed 40 sentences randomly selected from a set of 75. There were 20 sentences (trials) each for one-finger and two-finger typing, with each participant typing in both conditions (order was counter-balanced). No participant was given the same sentence twice.

4.3.3 Materials

Smartphones with a 4.7–5.5-inch touchscreen form the mainstream of the current market [79]. Among those devices, we chose the Samsung Galaxy S6 smartphone (1440 × 2560, 577 ppi) with a screen size of 5.1 inches. We developed a custom typing application for collecting key-pressing data and permitting the synchronization of data sources. The typing application is shown in Figure 4.2. Its two main views were the calibration view and the typing view, the former used for synchronizing data sources at the start of the task block and the latter for the transcription tasks themselves. The keyboard in the application had a standard Finnish Qwerty layout (key height: 10.06 mm). The participants transcribed relatively simple, memorable everyday sentences selected from the Enron Mobile Email Database [24, 80]. Seventy-five sentences were translated into Finnish by a native speaker and checked by one of the authors. All sentences were stripped of special characters and punctuation, and everything was in lowercase. The mean sentence length was 20 characters (SD = 4).

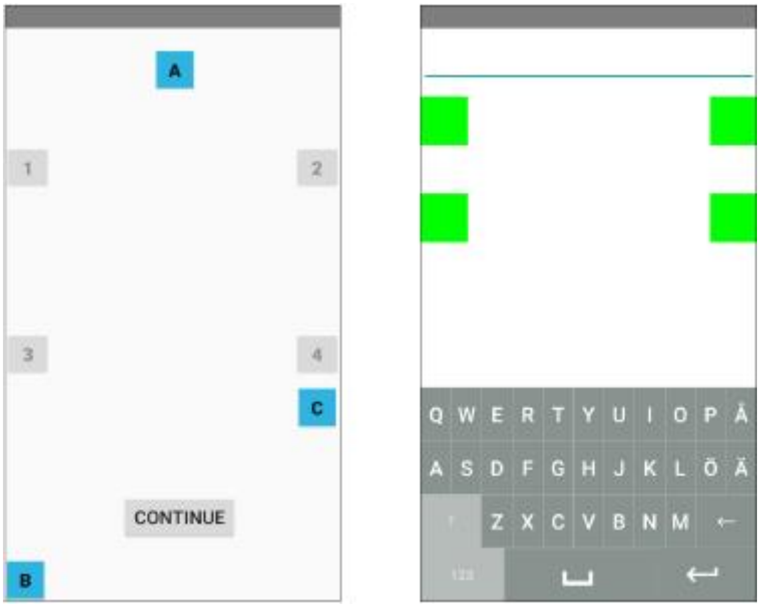


Figure 4.2 A view for calibrating the eye tracker (left) and the user interface of the typing task (right). The green boxes are for eye-tracking purposes.

Figure 4.3 Posture for holding the device and sitting during the experiment. Shown are grips for one-finger and two-finger typing. The block above the device is for tracking the phone position. [Photographed by the author at Aalto University, August 2018.]

4.3.4 Procedure

Firstly, participants were told that the purpose of the study was to analyze the movement of the eyes and fingers in smartphone typing, and they filled in a background questionnaire. During the experiment, they sat in a chair at an adjustable-height table with the smartphone freely in their hands, which were resting on the table (Figure 4.3). They were then given five minutes to practice and become familiar with the typing interface. After three-point calibration for eye-tracking glasses, participants were asked to press four buttons on the screen, marked with the numbers 1 to 4, in ascending order, for synchronization between the motion tracker and smartphone. Each task trial consisted of one sentence, which was given to the participant aurally via a speech synthesizer to avoid unnecessary eye movement during the experiment. The participants were asked to repeat the sentence aloud to confirm that it was heard correctly and to strengthen the memory of the sentence, after which they could start typing. They were asked to type as quickly as possible and not leave errors in the final sentence submitted. As the task block dictated, the participant used either two thumbs (two-finger condition) or the index finger of the dominant hand (one-finger condition). For two-finger typing, participants were asked to hold the device in both hands and perform typing with two thumbs. For one-finger typing, they were asked to type with only the index finger of their dominant hand and hold the smartphone on the other hand. During the experiment, we suggested that the participants rest their arms on the table and try to keep the same posture throughout the experiment block [81]. However, there were no physical constraints to movement, and the subjects reported no discomfort. Error correction could be performed via a backspace button, with no other means provided, such as moving the cursor by touching the typed text. The trial time for one sentence was calculated as the time from the first keypress to pressing Enter, keypress being defined as the moment of a key-down log event.

4.3.5 Data Collection and Preprocessing

We collected three types of data: eye movement, finger motion, and keypresses. For eye movements, we used SMI model 2W A eye-tracking glasses (60 Hz at 30 FPS). The glasses had infrared cameras tracking eye movements and a forward field camera to record the screen of the mobile device held in the hands. Participants with corrected vision had corresponding corrective lenses attached. The three-point calibration was done via the calibration screen (on the left in Figure 4.2), with the participant asked to focus on the blue rectangles one at a time. In the experiment proper, the green rectangles (in the right pane of Figure 4.2) were used to transform the eye-tracking coordinates into device screen coordinates.

To track finger movement, we used an OptiTrack Prime 13 motion-capture system that provides 3D precision of up to 0.2 mm at proximity. In one-finger typing, a reflective marker was attached to the top-middle part of the nail of the index finger of the dominant hand; in two two-thumb typing, one was attached to each thumb. The system was calibrated at the start of the block, with the same calibration screen as for the eye-tracking device (Figure 4.2); the participants were asked to type the numbered blocks in order. For turning the finger position into device coordinates, four reflective markers were placed above the smartphone, in a holder (Figure 4.3).

We checked all data manually and excluded three participants because of the loss of fixation data (resulting in N = 30). From the remaining participants' data, 244 trials out of the 1,199 were excluded from eye-movement-related analyses due to data corruption (i.e., fixation data were present for less than 90% of the trial). The loss was not correlated with sentence length (M = 20.44 words before data removal, M = 20.46 after). Loss of some motion-tracking data led to 45 further trials being excluded from finger-movement-related analyses (same criterion; also no change in sentence length). Finger tracking data were validated by confirming that the lowest local points of the finger(s) coincided with the pressing of keys in the device log. We extracted the coordinates from the raw data on finger and eye movements and converted them into a common coordinate system for the smartphone screen. In the data, the upper-left corner of the screen is the origin (0,0,0), with x-axis values increasing toward the right of the device and y values from top to bottom. The distance from the screen facing upward is the positive z value. The unit in a datum refers to one pixel of the smartphone screen. The motion-tracking system labeled and tracked each marker during the experiment. In the two-thumb typing condition, in cases where the tracker confused the fingers with each other due to their proximity, we checked and corrected the data manually.

4.3.6 Metrics

We followed the guidance for typing performance metrics [82]. For eye and finger movement data, we compare the metrics to the previous eye-and-finger-tracking study of physical keyboard typing [24]. The metrics used can be summarized thus:

- Inter-key interval (IKI) [24]: the time between two subsequent keypresses.
- Words per minute (WPM) [24]: the number of standard words (every five characters in the final input text) divided by the time spent on typing.
- Backspace [63]: the number of Backspace presses during typing of a sentence.
- Uncorrected error rate [82]: non-corrected incorrect keystrokes as a percentage of the sum of incorrect (whether fixed later or not) and correct keystrokes.
- Corrected error rate [82]: incorrect but rectified keystrokes as a percentage of the above sum.
- Immediate error correction [83]: the frequency of error correction in which the user immediately identified and corrected an error with a subsequent Backspace press.
- Delayed error correction [83]: the frequency of error correction wherein the user tried to correct previously missed errors in the middle of the input stream.

- Chunk length: the average length of a chunk during typing of each sentence. In typing, "chunking" refers to splitting the sentence into smaller pieces to manage working memory load [84]. We identified the border of a chunk when a clear increase in IKI is observed [85, 86]. The difference between neighboring IKIs is denoted as IKI difference (IKID). If the difference between neighboring IKIs is greater than the average IKID for the sentence, the key-pressing moment is considered to be a chunk border.

- Gaze shift: the average number of glances away from the keyboard area into the text area during typing of a sentence. The areas are defined as either the text area or the keyboard, both extended by 1.40 cm to all directions to account for foveal vision and possible slight drift in the eye-tracking data. Gaze shift has previously been measured as the number of gaze shifts from the monitor to the keyboard, reflecting most of the attention being put on the monitor [24]. We measured it in the opposite way, assuming that most of the attention would be on the touchscreen keyboard.

- Time ratio for gaze on the keyboard: the percentage of the time spent glancing at the keyboard. This is obtained by dividing the duration of gazing at the keyboard area by the total trial time [24].

- Entropy [24]: how consistently a key is pressed by the same finger in the two-finger typing condition. For each key k, given a frequency distribution over the two fingers, we compute the entropy as $H_k = -\sum_{f \in Fingers} p_f \log_2(p_f)$, where p_f is the probability of finger f pressing key k. The average entropy of a finger-to-key mapping is then computed as a sum over the entropy of each key weighted by the frequency of the corresponding letter. If a given key is always pressed by the same finger and this is true for all keys, the entropy is 0. To represent the finger–key mapping graphically, we show the distribution of touchpoints, using different colors for different fingers. Heatmaps were created on the background of a keyboard screenshot with layers of density plotted via the seaborn.kdeplot tool.

- Keys per finger [24]: the number of keys controlled by each thumb in the two-finger typing condition.

- Finger path: the distance that a finger has traveled during typing of a sentence.

- Distance to the next key [24]: at the moment of the current key-pressing, the average distance between the next target key and the finger for pressing that key.

- Finger alternation [24]: the percentage of bigrams entered with finger alternation.

- Same finger bigram [24]: the percentage of bigrams entered with the same finger.

- Letter repetition [24]: the percentage of pressed keys that are the same as the previous key.

Statistical tests were carried out using the Wilcoxon signed-rank test with $\alpha = 0.05$. Correlations between factors were calculated via Linear mixed-effects models with the lme4 package for R. Below, we report standardized βs as the correlation metric, noting any control variables that were used. In addition, all models had the task condition (number of fingers used) as a fixed effect, and subject and sentence-level as random effects [87]. The p-values for β estimates were calculated via Satterthwaite approximation to degrees of freedom.

4.4 Results

We collected, in total, 31,988 keypresses from the 30 participants (16,593 in the two-finger and 15,395 in the one-finger condition). Table 4.1 summarizes our main findings, aggregated first at the subject level and then on the grand condition level.

Table 4.1 Overview of the results for the one-finger and two-finger typing conditions. Differences between the conditions were tested using the Wilcoxon Signed-rank test (df = 29), with effect size d computed as Cohen's d value. [1]Not including consecutive clicks on the same key. *) $p < 0.05$ **) $p < 0.01$, ***) $p < 0.001$.

		Two-finger		One-finger		Wilcoxon test	
	Measure	M	SD	M	SD	W (29)	d
Performance	IKI (ms)	266.81	63.56	380.94	50.95	82***	-1.98
	WPM	39.33	10.3	27.19	3.61	779 ***	1.57
	Backspace	3.58	2.8	2.61	1.81	575.5	(0.41)
	Uncorrected error rate (%)	0.6	0.87	0.56	0.71	468.5	(0.05)
	Corrected error rate (%)	12.23	7.29	9.38	5.75	586*	0.43

		Two-finger		One-finger		Wilcoxon test	
	Measure	M	SD	M	SD	W (29)	d
	Immediate error correction	0.41	0.38	0.40	0.26	413	(0.03)
	Delayed error correction	0.93	0.83	0.63	0.47	569.5	(0.44)
	Chunk length	4.43	0.53	3.98	0.41	689***	0.94
Eye gaze	Number of fixations	18.79	8.05	24.04	4.56	192***	-0.81
	Fixation duration	315.27	67.65	303.99	45.72	454	(0.2)
	Saccade length (cm)	3.37	0.75	3.58	0.68	339	(-0.29)
	Gaze shift	3.4	2.31	3.91	1.5	269*	-0.26
	Time ratio for gaze on keyboard	0.6	0.16	0.7	0.14	263**	-0.69
Finger movement	Entropy	0.07	0.04	0	0	-	-
	Keys per finger left	4	0.68	12.5	0.55	0***	-13.75
	Keys per finger right	9.24	0.74			0***	-5
	Finger path left (cm)	23.06	2.82	25.29	1.33	179***	-1.02
	Finger path right (cm)	26.16	2.11			702***	0.49
	Dist. to next key[1] (cm)	1.2	0.15	2.3	0.1	0***	-8.78
	Finger alternation (%)	39.91	4.59	0	0	-	-
	- IKI (ms)	243.10	73.63	-	-	-	-
	Same finger bigram (%)	60.09	4.59	100	0	0***	-12.30
	- IKI (ms)	289.43	62.66	364.81	48.14	151***	-1.35
	Letter repetition (%)	11.59	3.1	10.17	2.37	591*	0.51
	- IKI (ms)	177.78	23.81	182.15	25.25	417	(-0.18)

4.4.1 *Typing Performance*

As expected, we found statistically significant differences in typing performance between two-finger and one-finger typing (all test results are in Table 4.1). Users made more errors when typing with two fingers than with one (corrected error rate M = 12.23% vs. M = 9.38%), although they corrected most of these before submitting the final sentence (the uncorrected error rates were M = 0.6% and M = 0.56% respectively). When using two fingers, participants were faster at typing (visible in lower IKI and higher WPM values), used longer chunks (M = 4.43 vs. M = 3.98), and made fewer gaze shifts between the keyboard and the text-entry area. The chunk sizes identified here are in line with the morphology of the Finnish language [88]. The comparisons between one-finger and two-finger typing are consistent with the previously observed error rates of 10.80% and 8.17%, and with typing speeds of 50.03 WPM and 36.34 WPM for two-finger and one-finger typing, respectively (the study cited did not include error correction) [3]. Participants exhibited more delayed error corrections than immediate error corrections, both in one-finger and in two-finger typing, meaning that most errors were detected only after typing of further characters.

We also correlated typing performance with background factors. Younger users were more likely to type more quickly and made fewer gaze shifts [63]. WPM values were negatively correlated with age ($\beta = -0.27$, $b = -0.64$, $p < .001$) while gaze shift had a positive correlation with age ($\beta = 0.33$, $p < .01$). We did not observe a correlation between error rate and age.

4.4.2 *Finger Movement*

Global finger movement. Global finger movement is the length of the total travel path of a finger or fingers in a sentence. We found significant differences in path length between both left- and right-hand data from two-finger typing and data for the dominant hand in one-finger typing. The average finger path in one-finger use (M = 25.29 cm) is shorter than the sum of the finger paths in two-finger use (M = 49.22 cm). Two fingers together travel more during typing than one finger because each is free to move while the other finger is typing. As we show below, this is likely to be related to the preparation of the free finger.

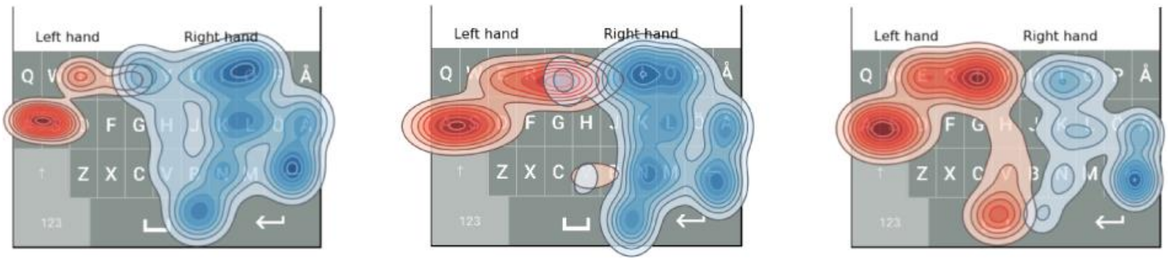


Figure 4.4 Heatmap showing finger-to-key mapping in two-thumb touch data of three participants (all sentences aggregated). The left thumb is shown in red, the right in blue. These patterns are representative of a tendency we found in the data for the right thumb to cover more keys than the left. The right hand was the dominant hand for most of the participants, but the same pattern was also observed for the left-handed participants.

Finger-to-key mapping. Work-based on logging data has assumed that the left and right thumb split the keyboard area [20]. We revisited that assumption in light of the motion-tracking data. We found that, overall, the right thumb (M = 9.24) is in charge of more keys than the left (M = 4.00). The right hand covers a larger area during typing than the left does, as Figure 4.4 shows. We observed no significant effect of finger-to-key mapping entropy on typing speed or error rate, meaning that this choice of strategy did not influence the participants' performance much. The reason could be that, since visual attention is needed for guiding the fingers, the finger that presses the next key can be selected opportunistically.

Finger preparation. Finger preparation is an index of how much a finger moves in advance even before its "turn". It is measured as the distance of a finger from the key that it will press at the moment when the previous key is being pressed. We observed shorter preparation distance for two-thumb typing (M = 1.20 cm) than for one-finger typing (M = 2.30 cm). Analyzing this further, we found a negative correlation between distance to the next key and WPM, $\beta = -0.41$, $p < .001$, even when controlling for the true distance between subsequent keys (in the two-finger condition, this refers to subsequent keys pressed by the same finger). As is visible in Figure 4.5, the effect is similar between the one-finger and two-finger conditions, although, understandably, the latter condition allows more flexibility for preparing the finger that is not currently typing. In one-finger typing, the keypresses are executed with the finger in a sequential manner, so the distance between the finger and the next key is approximately equal to the inter-key distance. However, in the case of two-finger typing, users are free to decide and can control their fingers for parallel input with the two fingers. As one finger is clicking on a key, the other finger is already activated for aiming for the next key. The parallel control visibly increased the typing speed in the two-finger condition.

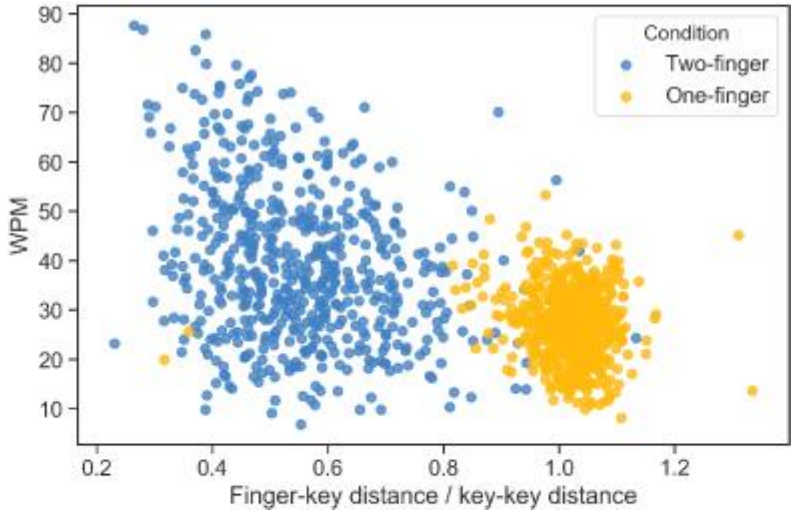


Figure 4.5 The impact of finger preparation on WPM by the task condition. The x-axis shows the distance of the finger from its next key, divided by the distance of the current and the next key (pressed by the same finger).

Finger alternation. Confirming previous findings based on log data [20, 63], we found a benefit for finger alternation. We observed a lower IKI in alternating (IKI = 243.10ms) as opposed to using a single finger (IKI = 289.43ms). This benefit notwithstanding, it was more common to continue using the same finger: 60.09% of bigrams were typed with one finger instead of two. Completing a bigram with one finger was faster in two-thumb typing than in one-finger typing (IKI = 364.81ms). This can be attributed to the longer average travel distance when one finger is used in typing. Figure 4.6 shows the IKI distribution between types of bigrams for the same finger and alternating fingers in two-thumb typing, the same finger in one-finger typing, and a repeated letter in both two- and one-finger typing.

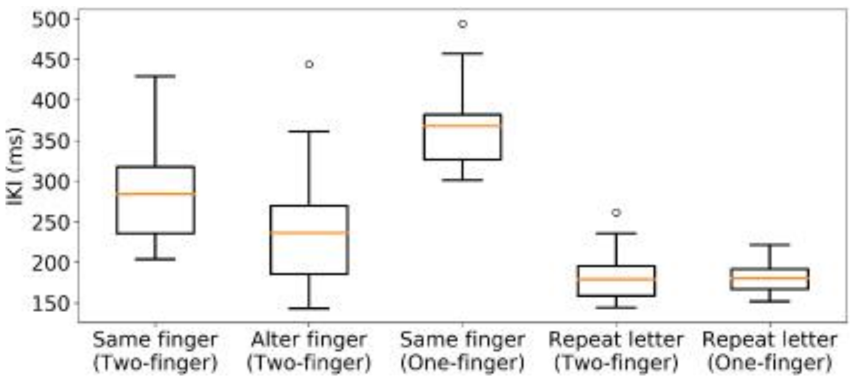


Figure 4.6 The impact of finger alternation for inter-key-intervals between different bigram types. Outlier data points are marked as circles beyond the caps of the boxes.

4.4.3 Eye-hand Coordination

Eye-hand distance. We examined the distance between the fixation point and the finger in one-finger typing (this cannot be unambiguously computed for the two-thumb case), so as to understand whether a closer eye-hand coupling can lead to better performance. We found a significant positive correlation between the average eye-hand distance and the average corrected error rate per sentence ($\beta = 0.32$, $p < .01$), controlling for the number of backspaces per sentence. This illustrates that typing errors are correlated with more visual attention to the text area, which results in looser eye-hand coupling. For typing speed, as expected, we found a negative correlation between average eye-hand distance and average WPM per sentence (Figure 4.7). However, this result may be explained by slower typists having to look at the text display more, which manifests itself here in long eye-hand distances. Therefore, we looked at eye-hand following specifically when both are operating in the keyboard area.

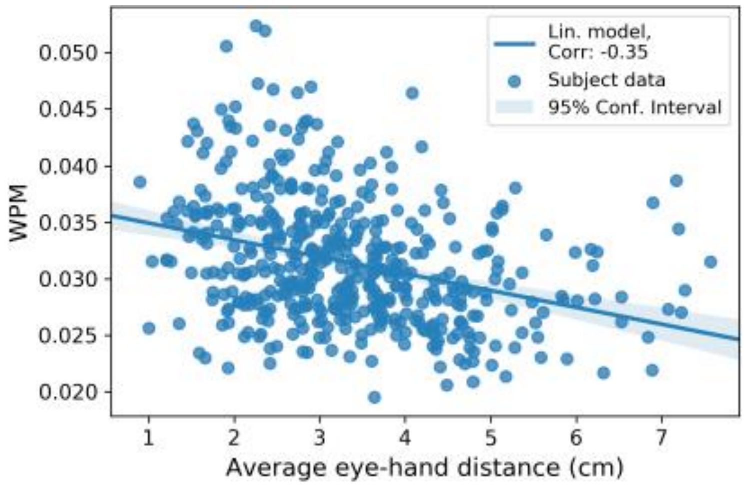


Figure 4.7 Eye-hand distance and average IKI per sentence.

Eye-hand following. When pressing a key or searching for one, the finger may follow the eye. To look at eye-hand following behavior, we extracted finger and eye movement paths in one-finger typing where the eyes stay in the keyboard area. We examined the dissimilarity between the finger movement and eye movement path by means of the Partial Curve Mapping (PCM) method, which uses a combination of arc length and area to determine the similarity between curves [89, 90]. We found a positive correlation between WPM and dissimilarity, $\beta = 0.16$, $p < .001$. If a user types more quickly, there is less similarity. One explanation might be that fast typists have less need to guide their fingers with the eyes and so retain global supervisory control over the keyboard while trusting in the accuracy of their fingers. Also, we found a negative correlation between time ratio for looking at the keyboard and dissimilarity ($\beta = -0.2$, $p < .001$), indicating, as expected, that the more the gaze is on the keyboard, the greater the similarity between the finger path and eye movement path.

A more detailed investigation of eye and hand movement might help to explain this finding. Figure 4.8 shows the distance of the eye and the finger from the next key that is typed, taken from two partial example sentences from two participants (one-finger typing). Glances at the text-entry area are visible as large distances between the eye and the target key. In both sentences, the finger and the eye move simultaneously toward the target key. The finger moves rapidly at first, and it slows down near the target for the final "peck". However, one can see a subtle difference between these participants: in the lower pane, the eye quickly finds the target key, after which it starts moving away from it even before the finger can peck it. Similarly, when backspacing, the upper-pane participant uses the eyes to locate Backspace, whereas the lower-pane one looks at the text-entry area after having already visually located Backspace. It is possible that participants who trust their pointing accuracy more can free their vision for other tasks than guiding the finger, producing both faster typing and greater dissimilarity between the eye and finger paths.

4.4.4 Proofreading and Error Correction

Gaze-shifting. We defined a gaze shift as a glance from the keyboard at the text area, and we take a gaze shift to indicate either (preemptive) proofreading or error-correction activity: a glance is initiated to check the typed text for errors, attend it for possible new errors, or to control backspacing when an error has been found. Again, in our study, the only way for participants to correct errors was by using the Backspace button. The left pane in Figure 4.9 illustrates the average number of gaze shifts with one-finger and two-finger typing, and for both sentences that contained error correction and those that did not. In sentences with typing errors (and the subsequent error correction), our participants shifted gaze between the text-entry area and the keyboard more than in error-free sentences. The pattern is identical between the typing conditions, although the one-finger condition displayed slightly more gaze shifts.

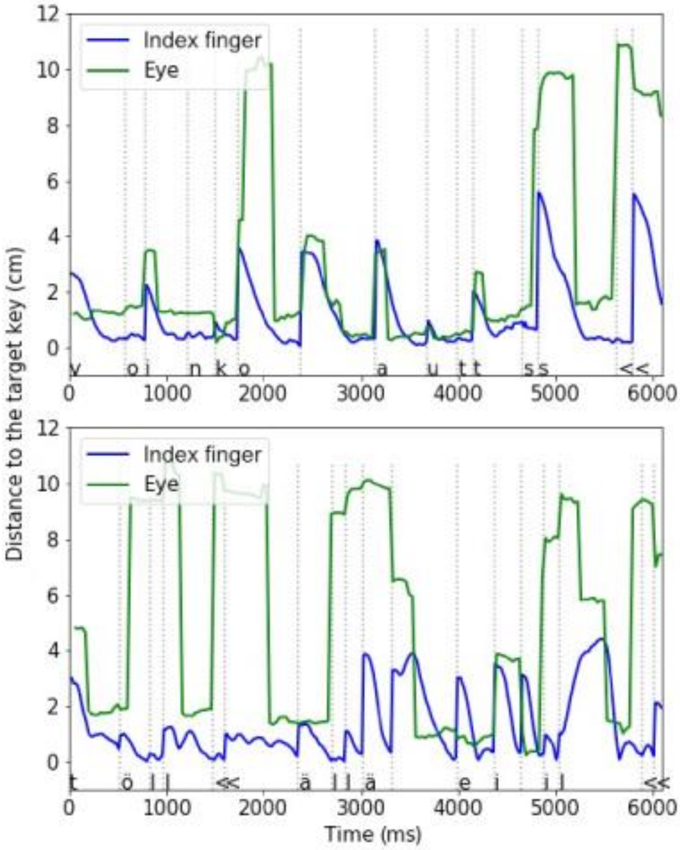


Figure 4.8 Key-by-key distance of the eyes and the finger in two partial sentences (truncated to about 6 seconds of typing; '<' refers to Backspace). Note the different sentences in these examples.

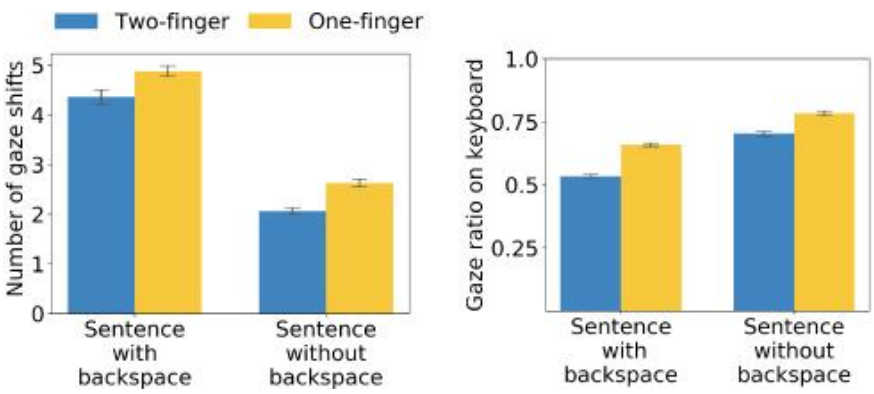


Figure 4.9 Number of gaze shifts to the text area (left) and ratio of gaze spent looking at the keyboard. Error bars are standard errors (right).

More total time was used for looking at the keyboard in one-finger than in two-finger typing, and more was used for sentences without backspacing (right pane in Figure 4.9). This means that glancing behavior is more erratic under one-finger typing, with gaze shifts as well as less relative time spent glancing at the text area. Analyzing the correlation between the time ratio for keyboard glances and the corrected error rate, when controlling for uncorrected error rate, we observed a negative $\beta = -0.35$, $p < .001$. Similarly, the number of glances at the text-entry area correlates with the number of corrected errors, $\beta = 0.58$, $p < .001$. Nevertheless, there were still, on average, 2.4 glances into the text-entry area for sentences that contained no error correction.

To investigate the impact of proofreading activity on typing performance further, we analyzed the correlation between WPM and gaze shifts, observing $\beta = -0.51$, $p < .001$. This negative correlation remained even after controlling for the amount of error correction the participants did (although the corrected estimate was smaller, $\beta = -0.17$, $p < .001$). Even for the subset of the data with only sentences containing no backspacing or uncorrected errors, we observed a negative correlation $\beta = -0.16$, $p < .001$. These findings mean that irrespective of the number of errors made and corrected, typing performance is negatively correlated with gaze shifting between the keyboard and the text entry area. Reflecting on the same phenomenon, the percentage of sentence-typing time for which the eyes were on the keyboard had a small but statistically significant correlation with WPM, $\beta = 0.11$, $p < .001$. Typists who focus more on the keyboard can reach higher text-entry rates. This focus may reflect a typist's level of confidence in not having made typing errors.

Correction of errors. We looked at two types of error correction: *immediate error correction* refers to when the user immediately identifies an error and corrects it with a subsequent Backspace press; *delayed error correction* occurs when the user attempts to correct an error in the middle of the input stream that was missed or overlooked, via multiple Backspace presses. To investigate the latter error type further, we split consecutive Backspace presses into the first press, intermediate backspacing, and the final press of Backspace.

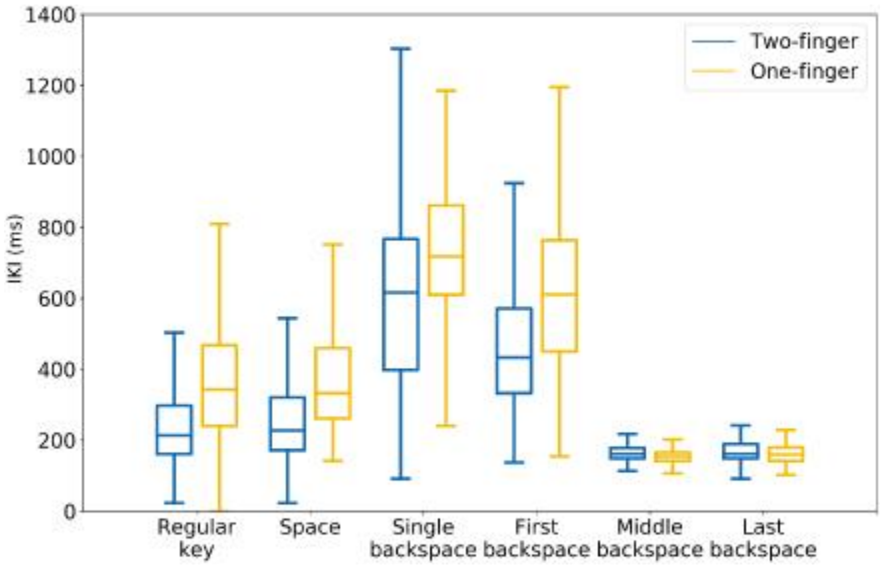


Figure 4.10 Typing interval for various types of key pressing.

As shown in Figure 4.10, it took more time for the participants to press Backspace a single time or make the first of multiple backspaces, relative to an average keypress. For intermediate presses during a run of backspaces and for the final Backspace press, the average time consumed was much lower than that for an average keypress. The average time used for a single backspace was lower in typing with two fingers than in typing with one finger.

We calculated the frequency of various types of error-correction behavior across conditions. For two-thumb typing, more delayed error corrections ($M = 0.93$) were observed than immediate ones ($M = 0.41$). Finally, we investigated the connection between typing errors on typing speed. Since errors and the backspacing that erases them (along with any correctly typed text between the error and the correction) do not contribute to typing the sentence, we expected a higher error count to contribute to smaller WPM. Controlling for uncorrected error rate, we indeed found a clear negative correlation between backspace count and WPM, $\beta = -0.61$, $p < .001$. Further, controlling for glances at the text entry area, the effect remained at $\beta = -0.49$, $p < .001$.

4.5 Discussion

A rich dataset was collected to deepen our understanding of how the fingers and eyes move in typing with mobile devices. The main finding over prior work is that movement strategies in mobile typing are strongly affected by competition for visual attention. Whereas with physical keyboards, a skilled typist can keep his or her attention on the text display, where it is needed for detecting errors [25, 26], in mobile typing, the need to guide finger motion competes for attention. Since one cannot monitor the keyboard and the text display at the same time, even though the mobile device is small, a strategy must be selected that determines which to give attention to and when. A good strategy must strike a compromise between the cost of not correcting errors early and the time lost in glancing at the text display when the fingers cannot be guided. Further, if the typist is not skilled with the keyboard, they need to conduct a costly visual search, which we did not need to consider in our analysis [56]. Conversely, it is possible that a very skilled typist has learned to control finger movement to the extent that most of the time the gaze can be kept on the text area. However, a more detailed analysis of very fast typists would be required to investigate this. Supervisory control in mobile typing is, hence, not just about the speed-accuracy tradeoff of finger movement; at its core is the deployment of gaze between the main regions of the application. While cost-benefit analyses have shown this in the case of intelligent text-entry methods [76], the general point has not been made before with support from data.

Understanding the competition for attention that goes on in typing helps us understand what makes typing fast vs. slow. It also makes important implications for smart typing aids, which in light of our results, should not compete for attention and require the learning of more complicated attention shift policies, like, for instance, word prediction lists do. We found that typing speed is positively correlated with the amount of attention on the touchscreen keyboard. The attention of a typical typist in our study was on the keyboard about 60% of the time, while the equivalent figure for a touch typist in a comparative study of physical keyboard typing was only 20% [24]. Also, the frequency of gaze shifts is much higher in mobile typing: 3.4 in our study, compared with 0.92 in the physical keyboard study. Similarities between mobile typing (our study) and typing on physical keyboards [24] include unequal division of labor between hands, the benefit of preparatory movements, and the negative effect of errors on typing speed. We further note that that study is similar to ours in key respects, including the task and the sample (our mean age 26, their 31 years).

The notion of resource competition can also help refine our understanding of the known superiority of two-thumb typing over one-finger typing [3, 63]. It has been attributed to the alternation between the lateral sides of the keyboard [20], and our results corroborate this.

Switching between sides is faster than moving a finger from one side to another. But we also found that two-thumb-typing users benefit from preparatory movements, moving a soon-to-press finger toward its next target, similar to a pattern found in typing on physical keyboards [24]. However, significantly more errors are made when typing with two fingers instead of one. We found that in two-thumb typing, there are more intervening keypresses between consecutive glances at the text display. Users notice the errors later. Hence, the large benefits of two-finger use (shorter travel distance, preparatory movements) outweigh the costs (more delay in detection of errors). Making users aware of possible errors earlier presents an interesting challenge for intelligent text entry methods.

We also observed a curious and previously unreported phenomenon in two-thumb typing: there is an unequal division of labor between the two lateral sides of the keyboard. Earlier models of two-thumb typing, based on log data rather than on direct observations of finger movement, have assumed equal distribution [20]. We found, in contrast, that the right hand does most of the work. This was the dominant hand for most of the participants, but the same pattern was true for the left-handed participants. The finger movement paths were significantly longer for the right (dominant-hand) finger than the left. While the monogram frequencies of Finnish might contribute to this, the effect has also been observed with non-Finnish typists using physical keyboards [24]. The unequal split between the hands could have implications for the customization and adaptation of keyboard layouts.

Our results can inform the development of predictive models. The strong role of gaze deployment we found is in stark contrast with some previous accounts that have framed mobile typing in terms of finger movement [35, 75]. What analyses based on Fitts' law miss is the significant challenge posed to visual attention in typing: how to juggle between the two areas of the display that need attention. Fitts' law conceals these intriguing and critical effects in the empirical parameters (a and b). At the same, the existing non-Fittsian models similarly fail to account for the parallelism of gaze and finger movements. The KLM model of Holleis et al. [57], its extension by Sarcar et al. [53], and the ACT-R model of Cao et al. [54] all assume that either the gaze or the finger is moving but not both. Finally, it is important to develop generative models that model how eye-hand strategies adapt – for example, to changing the probability of errors, to the number of fingers used, to the cost of error correction, and with time/experience. While these points have been made before [53, 56, 76], we currently lack a unified model. In light of our findings, such a model must be able to explicate the role of attention control. We propose that hierarchical reinforcement learning [91] is a potential candidate control principle to explain these gaze deployment strategies

4.6 Limitations and Future work

A few caveats must be taken into account when interpreting our findings. Firstly, our experiment was conducted in a quiet laboratory, with participants comfortably seated and resting their arms on the table. Mobile typing often takes place in dynamic environments, and there might be substantial differences in performance and strategies [92, 93]. The instrumentation we used in our study cannot be easily used in the wild, but it is possible to design laboratory interventions that emulate real-life circumstances, such as walking or multitasking (e.g., [94]). Second, our participants used a normal touchscreen Qwerty keyboard without intelligent text-entry aids, such as error correction or word prediction. As most smartphone users seem to be using intelligent aids [63], studying this phenomenon is important for future work.

Third, we asked the users to correct all errors, a practice followed in some but not all text-entry studies. On the one hand, this simplifies typing since one need not regulate which errors are to be left as-is and which not, but, on the other, this renders it more important for the user to check the text display because errors must not be left uncorrected. We believe that our main finding will not fundamentally change with the introduction of mobility, intelligent aids, or errors, but these could result in different attention-sharing strategies. We believe that these factors complicate the problem that cognition faces in typing simply because there are more tasks competing for visual attention [92]. Thus we expect even larger variability in movement strategies and, consequently, in typing performance. Fourth, our participants were relatively young adults with experience with technology, which is important to keep in mind since, for instance, older adults are generally slower at typing on smartphones [95]. A study with wider participant demographics is warranted for the future. The final caveat involves the language in the experiment. While we used a standard mobile corpus [80], the sentences were translated into Finnish, which has unique n-gram distributions and grammar. The keyboard layout we used also had two additional umlaut characters to the right. We do not expect the role of visual attention as reported here to differ greatly because of language, but the finger movement paths may vary between languages.

4.7 Summary

In this study, we report rich and detailed finger and eye movement data from mobile typing. We illustrate and discuss the role of visual attention in mobile typing, contrasting it to typing with physical keyboards. To facilitate further research on this topic, we have made the software and analysis scripts, along with all data and instructions on how to analyze it at https://userinterfaces.aalto.fi/how-we-type-mobile/.

CHAPTER 5

STUDY 3: UNDERSTANDING LEARNING AND USER BEHAVIOR ADAPTION ON RANDOMIZED KEYBOARDS

Typing on a mobile device requires users to have expertise like remembering and recalling the keyboard layout, controlling and coordinating multi-finger movements, key searching, and remembering the typing material. Although using soft keyboards on a daily basis, few people can claim themselves as expert typists. While motor control in typing is relatively well understood, much less attention has been paid to skilled use of visual attention, knowing where to look and when. This study investigates data from a controlled experiment where people typed with Qwerty and randomized keyboards, allowing us to trace the development of eye and finger movement strategies over time. We demonstrate how strategies, such as speed-accuracy trade-offs and gaze deployment between the different regions of a keyboard, are dependent on the amount of experience. The results suggest that, in addition to motor learning, the development of performance in mobile typing is attributable to the adaptation of visual attention and eye-hand coordination. In particular, the development of better location memory for the keyboard layout shapes the strategies.

5.1 Introduction

Although keyboards with Qwerty layout are widely applied for text input, users still encounter new keyboard designs that require learning before being mastered. During this learning process, users transfer from “hunt-and-peck” to a more skilled typist by building an internal representation of the task, including the memory of the keyboard layout. Memorization of the keyboard layouts helps to minimize the time needed for visual search [56], while at the same time, increases the certainty of finger touch operations on the flat screen. By repetitive practice, users can further build internal mappings between the layout and finger movement control [96]. Understanding how users develop their skills with new keyboard layouts can help researchers and designers discover more effective ways to improve interfaces of text input applications. Our theoretical framework for understanding changes that are associated with the transition from novice to skilled typists is that of adaptation to available resources [97]. We assume as users become familiar with a UI, they develop certain resources and skills, such as robust memory of the layout and ways to access key functionalities. At the same time, their behavior adapts to account for these changes. For instance, in touchscreen typing, a novice begins without a memory of the location of the keys, resulting in long intervals between typed keys due to time spent visually searching the keyboard [56, 98]. Our results demonstrate that the constraint of not having a memory of the UI is obvious not just in finger movements but also in other phases of typing, such as proofreading and error correction. This is an important reminder for touchscreen typing research, which has previously focused mainly on the motor performance of skilled typists. Visible from our results is the fact that, because novices have to spend time visually searching between key-presses, they emphasize accuracy over speed, and they are able, over time, to reduce the number of gaze shifts for proofreading as they become confident about their accuracy. As the users become more familiar and confident with the keyboard layout, they permit more errors due to the lower relative cost of having to retype characters, leading to even more proofreading activity as well. Finally, as users become more and more familiar with the keyboard layout, they type progressively faster, resulting in the introduction of more and more typing errors. However, these behaviors of a higher level of skill are not detrimental to their performance, as it permits very fast movements, minimizing time spent between keystrokes. Furthermore, as the eyes are not constantly needed to guide the movement of the fingers, a larger fraction out of the total typing time can be spent on proofreading.

In this study, we captured the details of eye and finger movements throughout the typing process with an experimental control on the participants' level of typing skill. By randomizing the keyboard layout, we created three levels of the learning process for the participants. The skilled performance was obtained using a conventional Qwerty layout, as all the participants are familiar with it. By assigning each participant their own once-randomized or statically randomized (SR) keyboard, we could observe the learning process and adaptations made throughout the trials of sentence typing. Finally, using a constantly or dynamically randomizing (DR) keyboard to shuffle the keys after each key-press, we were able to keep the participants at the novice level throughout this part of the experiment. By doing this, we not only compared the different levels of typing skills, but also observed the learning process for each new keyboard layout, from the perspective of eye and finger movement control.

The contribution of this work is presenting and explaining the learning process in the early stages of typing skill development on new keyboard layouts, based on detailed capture and analysis of eye and finger movement. We explain how users adapt their strategy of speed-accuracy trade-off and proofreading behavior based on the development of their internal model of the keyboard layout, and related skills such as key-searching and finger movement. Findings provide a reference for designers aiming for improved UIs with layout reform considering different levels of typing skill, as well as for researchers developing mental models for behavioral strategies prediction.

5.2 Related Work

Typing behavior on the touchscreen has been studied for decades from the perspective of performance like speed and error rate. Although such metrics can provide quick and easy-to-understand results for purposes like technique validation and comparison, they cannot indicate more detailed behavioral differences such as in movement control and attention shift. Recently, studies have started to emerge on analyzing and modeling user behaviors during typing, such as touchpoints [3, 35], eye movements [53, 56, 71], and finger movement [24, 99]. Here we introduce the related work on understanding typing-related behaviors on mobile touchscreen devices.

5.2.1 Typing Performance on Mobile Devices

Although it has been 30 years since the first smartphone was invented [100], typing speed on mobile touchscreen devices is still generally slower than on a physical keyboard [10, 63, 99]. While the average expert typing speed on a physical keyboard can reach 56 WPM (words per minute) [101], average typing speeds on smartphones are still around 30 - 40 WPM [63, 99]. Studies have explored the factors which affect typing speed, including the number of fingers used (one finger or two fingers) [3, 6, 15, 99], the size of the device [8], and the virtual keys [7], and different typing tasks (copy typing or memorized typing) [10]. Even though limitations on the key size present challenges for typing, some fast typists can still reach speeds of over 80 WPM [63].

Errors happen more frequently during typing on touchscreen devices (between 7% and 10.8% [3] compared with on physical keyboard (between 0.47% and 0.76% [24]. It is hard for users to ensure correct keystrokes without tactile feedback on the flat, featureless screen. To discuss detailed differences in error correction behavior, two metrics were proposed to quantify: 1) Immediate error correction, meaning that the users correct errors soon after they occur, with only one backspace and a correct keystroke; and 2) delayed error correction, indicating that the user realizes the error some moments after it happened. In such situations, users have to press multiple backspaces and then input the correct and deleted letters. It was found that typing on a mobile device with two thumbs resulted in more delayed error correction (0.93%) than when typing with one index finger (0.41%) [99].

The one direct reason for errors in typing is the failure of the finger to land inside the bounding box of the keys. Studies found that the touchpoints for keys are generally distributed below the center of the keys [3, 5], indicating that users prefer less occlusion of the button during key-pressing. To help users confirm the touchpoint locations, measures like showing the touchpoint with dots were presented and proved to be effective in reducing the error rate [5]. However, showing touchpoints also slowed down the typing speed by up to 5.2% [5] due to more attention being paid to the dots instead of to the typing task itself.

5.2.2 Learning in Typing

Debates on problems of Qwerty layout continued for decades, from the application of typewriters to the use of keyboards in a variety of scenarios such as mobile devices and virtual reality. Researchers proposed new keyboard layouts with different key arrangements [18, 21, 56], key shapes [1, 102], letter groupings [23, 53], even different input methods (e.g., input by gaze following [103]. Such designs and optimizations brought new possibilities for improving the user experience in typing. However, new technologies take time to be learned and mastered. Previous work on typing behavior with a stylus pushed the user back to ‘novice behaviors’ by randomizing the soft keyboard after each tap on the keys. These works found that the typing speed on random keyboards (around 5 WPM) is significantly slower than on a Qwerty keyboard (around 20 WPM) [104].

Skill acquisition proceeds through three main phases [105]. First, in the initial cognitive phase, individuals learn the underlying structure of the activity and develop strategies for the task. Novice typists in this phase usually type with a "hunt-and-peck" method while building the memory of the keyboard layout. In the second associative phase, elements that are necessary for successful task execution become integrated into sequences of actions. Typists type faster with the help of memory. Finally, in the autonomous phase, performance becomes more automatic and poses fewer demands on attentional resources. Typists gain higher expertise and become able to control the keyboard with a "touch-typing" method. Generally, it takes 90 hours for a user to transform from a novice typist to being relatively skilled typist on the physical keyboard (typewriter) [106]. Learning is required when the user is not familiar with the underlying structure of the activity or lacks fully developed strategies for the task [105, 107]. Even switching two key locations on a Qwerty keyboard can lead to an increase in key-search time of an experienced typist [56]. As the typing skill is largely affected by learning and practicing [107], understanding how users acquire and maintain typing skills is vital to the development of typing techniques.

As typing skill has been mastered pervasively among touchscreen users, to capture the learning process of typing, we pushed the users back to a novice state by randomizing the keyboard layout. In this way, the existing well-built memory of the keyboard layout is no longer able to serve. Other research methods of the same purpose can also be found as scrambling letter orders in words [108, 109] and adding noise masks on the stimuli of the target materials [109].

5.2.3 Eye and Finger Behavior on the Touchscreen

Operation on a touchscreen is largely guided by vision, as the flat screen is not capable of providing enough tactile feedback for a "blind" touch. However, frequently switching attention between screen areas reduces typing speed. In the study of typing on a split keyboard on the tablet, subjects were asked to use peripheral vision to guide their finger movements instead of an eyes-on manner. They found that using peripheral vision reduced the attention switch and led to a 28% faster typing speed (27 WPM) over the typical eyes-on typing mode [74]. As for typing on a smartphone, the keys are normally smaller than the fingertips due to the limitation of screen size. In such conditions, touch is guided by the eye more than half of the time (60% for two-finger typing, 70% for one-finger typing) to operate more accurately [99]. However, as proofreading is also needed during typing, users have to frequently shift their attention between the text input area and the keyboard. Such attention shift leads to temporal breaks of the current task and switches to another, which takes cognitive resources and time. Using two thumbs, users reduced the frequency of attention shift. However, the price is that more errors (especially with delayed error correction) happened during typing. Such findings indicated the importance of understanding eye and finger movement in the discussion of the speed-accuracy trade-off during typing.

In the learning of typing on a new keyboard layout, we believe that there exists a dynamically changed balance of attention shift, speed, and accuracy. In this study, we would like to explore the details in this process and compare it with the experienced and novice typing behavior. Findings will provide references for not only the keyboard designers, but also the users of soft keyboards on mobile devices.

5.3 Method

The experiment was conducted after study 2, with the same setup and participants. We designed the experiment for capturing the detailed eye and finger movements in a transcription task on the keyboards with different levels of randomization. The sentences participants typed were also selected from the same corpus, which contains representative everyday messages. The keyboard used was designed based on the standard Finnish Qwerty keyboard on mobile devices, without intelligent typing aids.

While study 2 reported one-finger and two-finger typing behavior on the Qwerty keyboard, here in study 3, we look at conditions of keyboard randomization, and compare the behavior with one-finger typing on the Qwerty keyboard. We are interested in the eye and finger movement adaption after a new keyboard layout was applied, and compare with the skilled and novice typing behavior. Participants were required to transcribe sentences with everyday words and correct errors if they happened. All the typing was done by using the index finger of their dominant hand (Figure 5.1, left). In order to ensure that the participant focused on typing itself without other distractions, we used the keyboard without any intelligent typing aids. The interface of the experimental application can be seen in Figure 5.1. The green boxes on the interface were for data collection purposes. We manually checked the gaze data after the experiment and confirmed that the boxes did not affect eye movement so much. We tracked the eye movement, finger movement, together with the typing log on the smartphone screen. All data were synchronized in time, and all positions were transformed into a unified coordinate system.

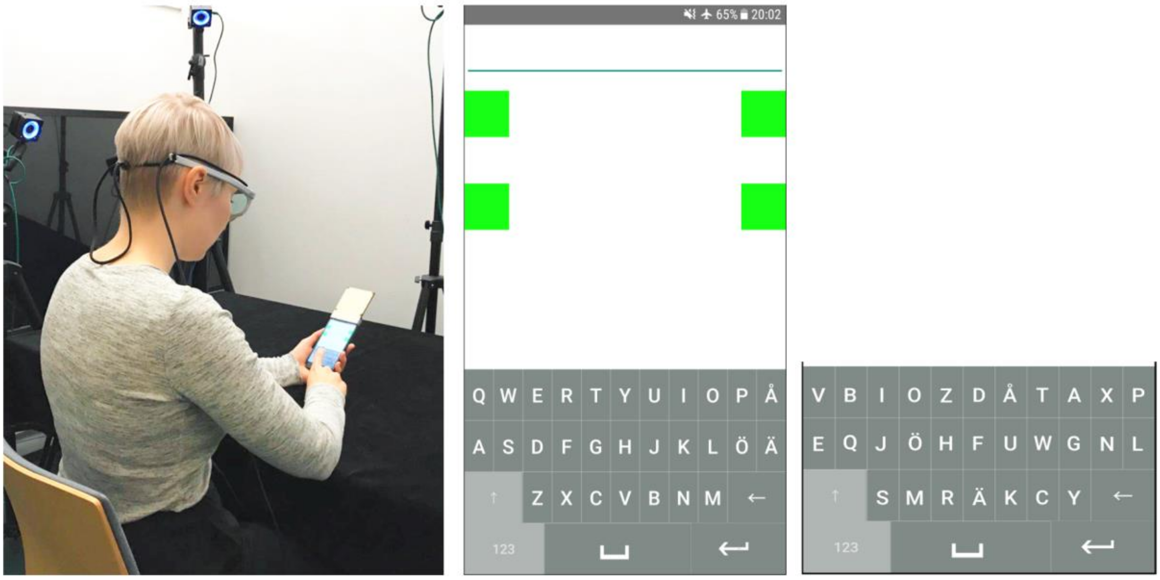


Figure 5.1 The experimental setup (left), typing application (middle), and randomized keyboard (right). [Photographed by the author at Aalto University, August 2018.]

5.3.1 *Participants*

The participant was the same as in study 2. The 33 subjects recurred were reduced to 30 (18 females; age range 18–45, M = 25.5, SD = 5.9) due to gaze-data loss (device error). Three of the participants were left-handed (1 female). All participants were native Finnish-speakers and had a normal or corrected vision (correction strength between -4 and +4). In this study, we observed a typing performance between 19.1–33.3 WPM for one-finger typing on a Qwerty keyboard. Each participant was compensated with two movie tickets (total worth about € 20) for their time.

5.3.2 *Experiment Design*

Each subject typed 20 sentences (trials) on the statically randomized (SR) keyboard. As typing on a dynamically randomized (DR) keyboard could lead to fatigue both physically and mentally, we reduced the number of trials to 15. We also compared the user behavior between typing on a Qwerty keyboard with one finger (data captured in study 2) and behavior on the SR and DR keyboards.

5.3.3 *Materials*

We used the Samsung Galaxy S6 smartphone (1440 × 2560, 577 ppi) with a screen size of 5.1 inches for the experiment. A custom typing application for collecting key-pressing data and the synchronization of data sources was developed. The typing application is shown in Figure 5.1 (middle). The Qwerty keyboard in this study refers to a standard Finnish Qwerty layout with a key height of 10.06 mm. The sentences used in the experiment were relatively simple, memorable everyday sentences selected from the Enron Mobile Email Database [24, 80]. Seventy-five sentences were translated into Finnish by a native speaker and checked by one of the authors. During the experiment, the target sentences were randomly selected from the seventy-five sentences. All sentences were stripped of special characters and punctuation, and everything was in lowercase. The mean sentence length was 20 characters (SD = 4).

5.4 Results

Totally 39,411 keypresses were captured in the typing log, in which 15,645 for Qwerty keyboard condition, 13,973 for statically randomized keyboard condition, and 9,793 for the dynamically randomized keyboard. In the following sections, we present the results and data analysis based on typing behaviors (i.e., typing speed, error rate, etc.) and eye and finger movement (i.e., gaze shift, finger travel distance, etc.).

5.4.1 *Typing Performance*

Trials with an uncorrected error rate higher than 25% were removed before analyzing the typing performance, following previous work [12, 16, 63]. The metrics are presented in Table 5.1. On average, typing on the dynamically randomized keyboard (DR) tended to be slow (6.99 WPM), which approximately equals to a quarter of the typing speed on the Qwerty keyboard (28.59 WPM). To test the effect of learning throughout the trials (Figure 5.2), we ran a linear mixed model analysis in R (version 3.5.3) using the lme4 package [110]. Significance was calculated using the lmerTest package [111], which applies Satterthwaite's method to estimate degrees of freedom and generate p-values for mixed models. Including participant as a random effect, we found that there are significant interaction effects of trial and condition on typing speed (i.e., WPM): Qwerty (beta = 0.89, $t = 28.44$, $p < .001$), SR (beta = -0.15, $t = -4.93$, $p < .001$), DR (beta = -0.85, $t = -19.19$, $p <$

.001). Then we filtered the data by condition, and found significant effects of trial for Qwerty (beta = 0.11, t = 2.37, p < .05) and SR (beta = 0.15, t = 6.95, p < .001). But there is no significant effect for DR (beta = 0.02, t = 1.91, p = .006).

Table 5.1 General typing performance for Qwerty, statically randomized (SR), and dynamically randomized (DR) keyboard.

	Qwerty	SR	DR
IKI	381.06 (68.80)	924.18 (284.40)	1787.61 (416.58)
WPM	28.59 (7.67)	13.55 (4.31)	6.99 (1.52)
KSPC	1.26 (0.37)	1.13 (0.30)	1.07 (0.15)
Corrected error rate (%)	9.44 (11.63)	4.63 (9.31)	2.77 (5.82)
Uncorrected error rate (%)	0.49 (1.65)	0.55 (1.97)	0.45 (1.57)
Immediate error correction	0.57 (0.50)	0.32 (0.47)	0.29 (0.46)
Delayed error correction	0.47 (0.96)	0.15 (0.53)	0.11 (0.41)

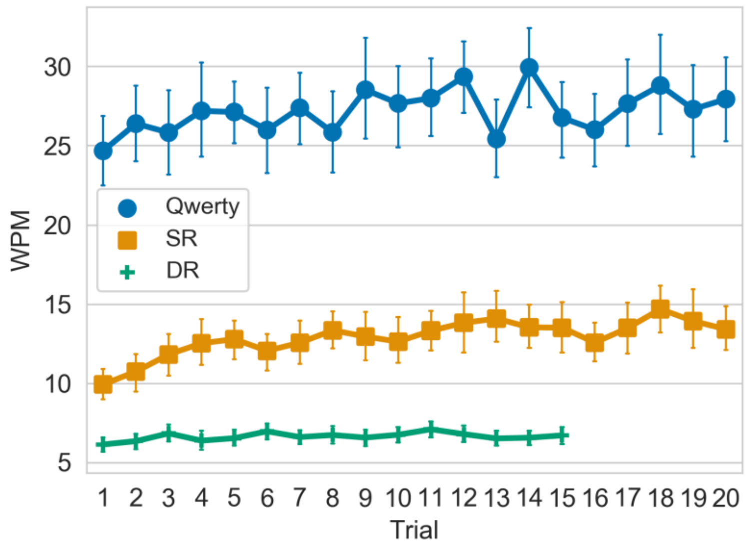


Figure 5.2 The change of average typing speed through trials.

As for the error correction behavior, we first measured the keystroke per character (KSPC), corrected and uncorrected error rate, following the measures presented in [112]. The total error rate, calculated as the sum of corrected and uncorrected error rates, was the highest on the Qwerty keyboard (9.93%) and the lowest on the DR keyboard (3.22%). While the same trend was also found in the corrected error rate, the uncorrected error rate showed the highest value on the SR keyboard (0.55%). Apart from the perspective of corrected and uncorrected errors, we also measured the number of immediate and delayed error corrections for each trial of sentence typing [99]. Immediate error correction refers to the errors immediately corrected with one backspace-pressing when it appears in the text input area. Delayed error correction refers to the situation in which the subject realizes the error a few keystrokes after it happens. Normally, the delayed error correction contains the deletion of multiple letters and the input of correct and deleted ones. More immediate error correction means that the users tend to be more sensitive about the content of the input. Controlling the number of backspaces, we found that participants conducted more immediate error corrections in DR ($M = 0.75$) compared with SR ($M = 0.56$) and Qwerty ($M = 0.45$). The effect of condition on immediate error correction (controlling the number of backspaces) was found to be significant, with both the SR and Qwerty differing from the DR ($\beta_{SR} = -0.17$, $SE = 0.04$, $t = -4.38$, $p < .001$, $\beta_{Qwerty} = -0.29$, $SE = 0.03$, $t = -8.29$, $p < .001$).

5.4.2 Eye Movement

Some of the gaze data could not be captured due to occasional occlusion or other technical problems. Before conducting gaze-related analysis, we dropped trials with gaze data captured less than 90% of the keystrokes, resulting in 376 trials dropped.

We measured gaze shift as the number of eye movements from the keyboard to the text input area during each trial of sentence typing (Figure 5.3, Gaze Shift). We found that gaze shift was the most frequent on DR keyboard ($M = 4.53$, $SD = 1.33$), followed by Qwerty ($M = 3.91$, $SD = 1.48$) and SR ($M = 2.86$, $SD = 1.33$). The effect of condition on gaze shift was found to be significant, with both the SR and Qwerty differing from the DR ($\beta_{SR} = -1.83$, $SE = 0.17$, $t = -10.79$, $p < .001$, $\beta_{Qwerty} = -0.79$, $SE = 0.17$, $t = -4.66$, $p < .001$). As error correction requires both proofreading and finger guiding, more error correction would increase the possibility of more gaze shift. Controlling for the error rate, we found that users conducted most gaze shift in DR ($M = 1.40$) compared with SR ($M = 0.55$) and Qwerty ($M = 0.39$).

In order to quantify the attention on the keyboard for key-searching and finger guiding, we measured the ratio of time that gaze stays on the keyboard area (Figure 5.3), Gaze Keyboard Ratio). We found that, the gaze ratio on the keyboard was the highest for typing on a SR keyboard ($M = 0.86$, $SD = 0.11$), followed by DR ($M = 0.85$, $SD = 0.10$) and Qwerty

(M = 0.70, SD = 0.14). The effect of conditions on the gaze ratio on the keyboard was significant between DR and Qwerty (β_{Qwerty} = -0.14, SE = 0.01, t = -14.34, p < .001). However, no significant difference was found between DR and SR (β_{SR} = 0.01, SE = 0.01, t = 1.12, p = 0.26). The gaze movement reflected that typing on an unfamiliar keyboard (i.e., SR and DR) increased the attention on the keyboard area.

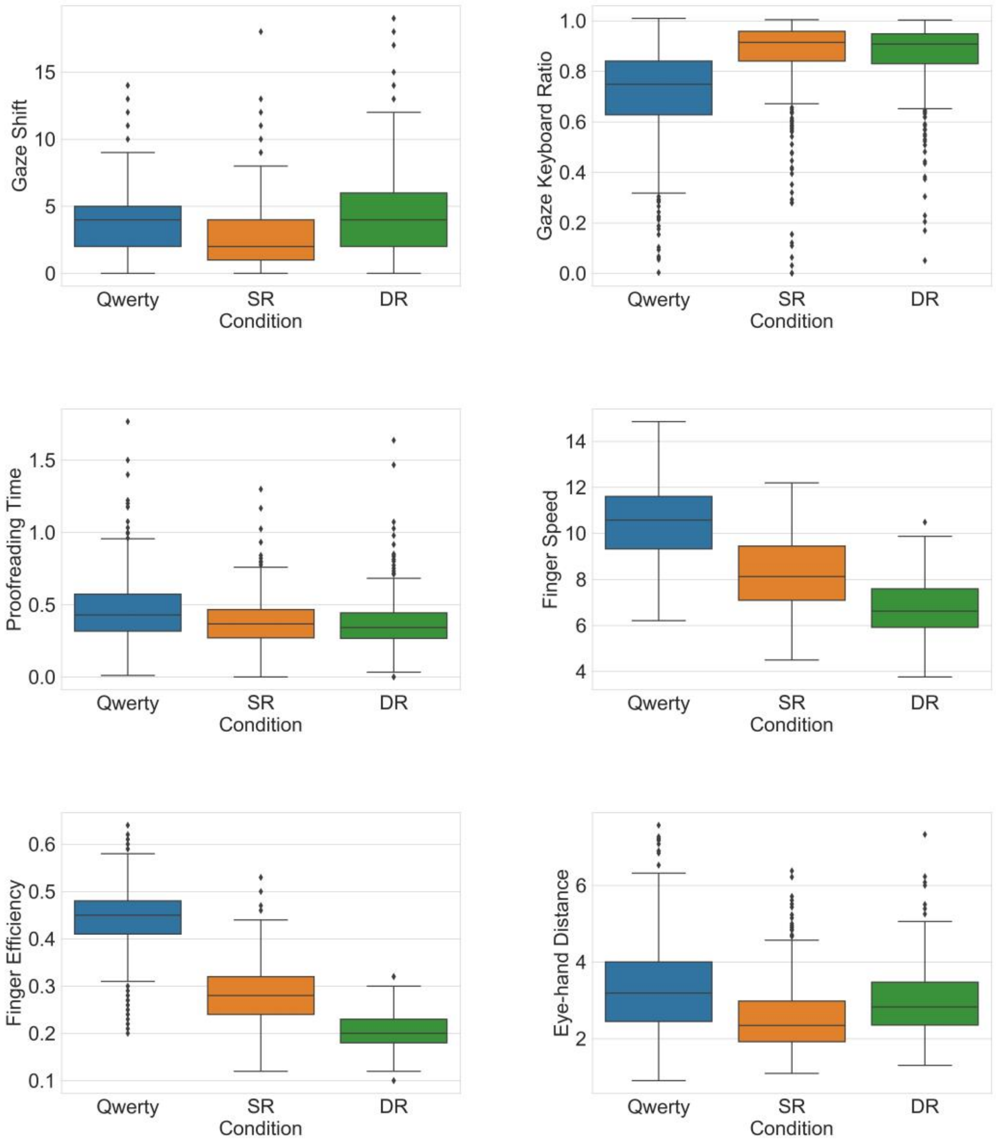


Figure 5.3 Box plots of eye and finger behavior.

Nevertheless, typing on a dynamically randomized keyboard without a fixed layout made users uncertain about their typed texts and conducted more proofreading. Proofreading time was measured in seconds as the duration in which the gaze stays in the text input area for each proofreading (Figure 5.3, Proofreading Time). Generally, subjects spend more time for each proofreading while typing on a Qwerty keyboard (M = 0.45, SD = 0.14), followed by similar duration of DR (M = 0.38, SD = 0.11) and SR (M = 0.38, SD = 0.09). The effect of condition on the proofreading was found to be significant between DR and Qwerty (β_{Qwerty} = 0.09, SE = 0.01, t = 6.30, p < .001). Similarly, to the gaze ratio on the keyboard, no significant difference was found between DR and SR (β_{SR} = 0.00, SE = 0.01, t = 0.34, p = 0.73).

5.4.3 Finger Movement

Detailed finger movement data were collected during the experiment, which enables us to measure finger behavior by using the real-time coordinates during typing. To ensure the data quality for analysis, we dropped trials with finger data captured for less than 90% of the keystrokes, resulting in 20 trials dropped. We measure average finger movement speed (centimeter per second) for each trial (Figure 5.3, Finger Speed), and found that in accordance with typing speed, finger moved the fastest on Qwerty keyboard (M = 10.44, SD = 1.57), followed by SR (M = 8.22, SD = 1.50) and DR (M = 6.77, SD = 1.25). The effect of condition on finger speed was found to be significant, with both the SR and Qwerty differing from the DR (β_{SR} = 1.5, SE = 0.06, t = 24.73, p < .001, β_{Qwerty} = 3.7, SE = 0.06, t = 62.53, p < .001).

As moving the finger across distance takes time, the ideal strategy is to keep a direct movement path between keys during typing. To see how efficiently the finger moved between keys, we report finger movement efficiency calculated as the inter-key distance divided by the finger movement path in the 3D space (Figure 5.3, Finger Efficiency). On average, users moved their fingers more efficiently while using a keyboard with Qwerty layout (M = 0.44, SD = 0.07), followed by SR (M = 0.28, SD = 0.07) and DR (M = 0.20, SD = 0.04). The effect of condition on finger movement efficiency was found to be significant, with both the SR and Qwerty differing from the DR (β_{SR} = 0.08, SE = 0.003, t = 28.06, p < .001, β_{Qwerty} = 0.24, SE = 0.003, t = 84.52, p < .001).

5.4.4 Eye-hand Relationship

We dropped 385 trials of eye and finger data before conducting the analysis, as they were either lacking data on eye or finger movement. We measured (in centimeter) the average distance between the gaze point and the projected finger coordinate on the touchscreen throughout the typing process of each trial (Figure 5.3, Eye-hand Distance). The highest distance was found with Qwerty keyboard (M = 3.33, SD = 1.22), followed by DR (M = 2.99, SD = 0.89) and SR (M = 2.53, SD = 0.89). The effect of condition on the distance between finger and gaze point was found to be significant, with both the SR and Qwerty differing from the DR ($\beta_{SR} = -0.49$, SE = 0.06, t = 5.29, p < .001, $\beta_{Qwerty} = 0.32$, SE = 0.06, t = 5.29, p < .001).

5.5 Discussion

Here we explain the process of typing skill development by going through the behavior adaption throughout trials during the experiment. We group the data into five sentences per trial block and explain the learning process by looking at the metrics of IKI, gaze keyboard ratio, finger efficiency, immediate error correction, gaze shift, gaze path per character, number of keys before proofreading, and total error rate (Figure 5.4). The metrics were first Z-score standardized within conditions to facilitate comparison of how they change between the trial blocks: in the charts, 0 refers to the average within-condition value, and one unit in the y axis is 1 SD of change.

We explain the behavioral difference by selectively interpreting metrics in Figure 5.4. First, looking at IKI, learning across all three conditions is visible with lower relative intervals by trial. However, as expected, this effect is largest for the SR condition and smallest for the DR. We can assume that even with participants who are familiar with Qwerty, there is some learning of the unfamiliar device used in the experiment. With SR, in addition to this learning, visible is the effect of learning the layout itself. Whereas for Qwerty, the overall improvement in IKI is a bit more than 0.2SD, for SR, it's more than 0.6S.

Gaze shift was not largely affected by practicing in Qwerty and SR (variation around 0.1 SD), while the number of gaze shifts per sentence is visually increasing throughout the trials of typing (0.3 SD). We assume that the reason for the trend is related to the fatigue of constant key-searching while typing on a dynamically randomized keyboard (DR). Subjects gradually lost the confidence of typing correctly and conducted more gaze shifts to the text input area for proofreading. A more efficient key searching can be seen on the decrease of gaze path per character in SR (around 0.4 SD), compared with an increase in DR.

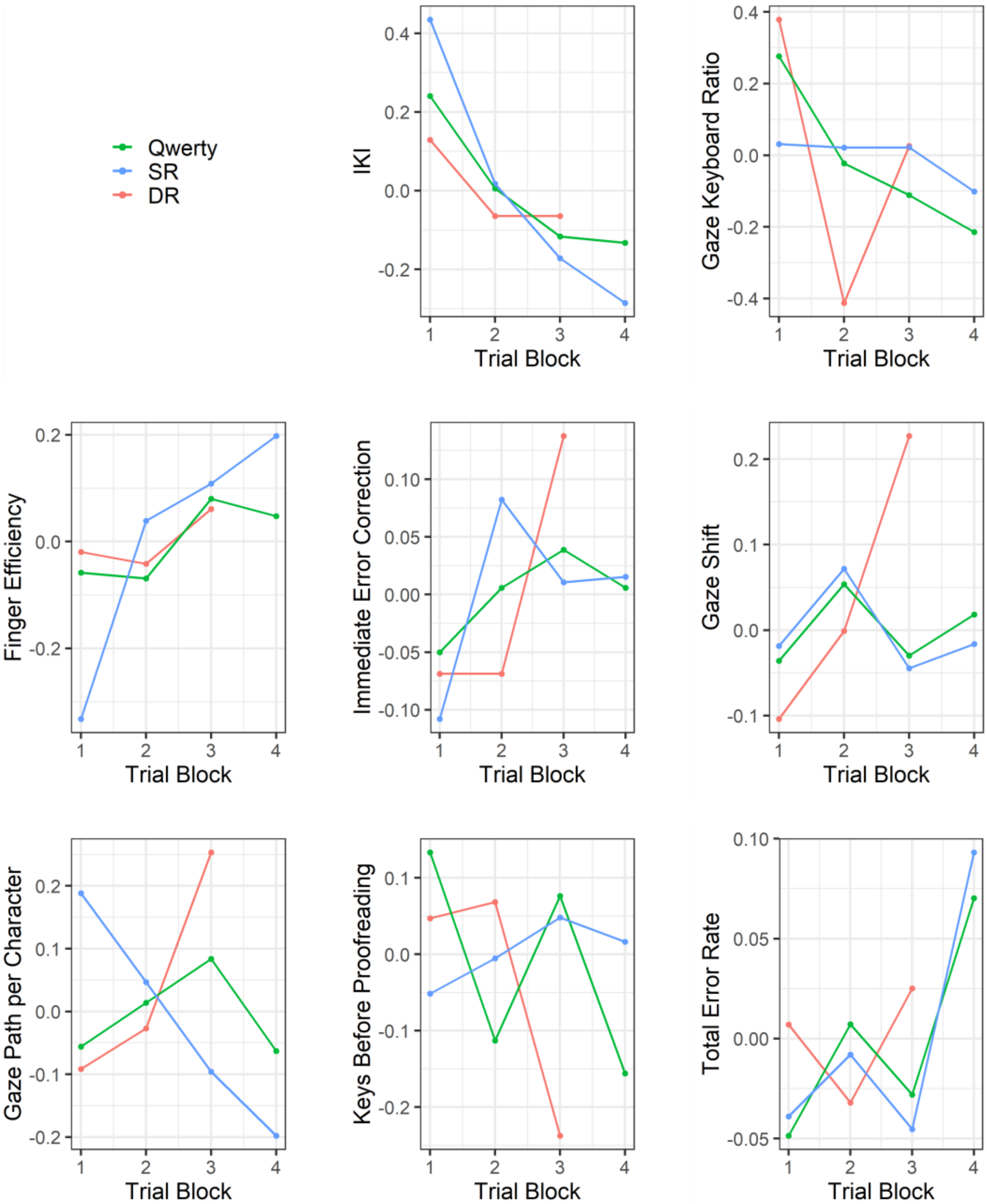


Figure 5.4 Change of eye, finger, and typing performance through time.

As for finger movement, an apparent increase of finger efficiency in SR can be observed throughout the trials of sentence typing (more than 0.5 SD), compared with variation around 0.2 SD for both Qwerty and DR. As the users gradually built the memory of the keyboard layout, their fingers moved more efficiently with better guidance. Next, we describe in chronological order how user behavior adapted throughout the development of typing skills.

5.5.1 *Development of Typing Skills*

When given a new keyboard layout, users can be regarded as a novice as there was no memory of the keyboard layout to guide their typing operations. They had to visually search for the keys before click, leading to relatively long inter-key intervals (Figure 5.4, IKI - SR - Trial Block 1) and longer gaze path traveled before each keystroke (Figure 5.4, Gaze Path per Character - SR - Trial Block 1). Under such constraints, the cost of each keystroke was the highest. To minimize the time cost for error correction, users tended to be extremely careful to make as few errors as possible, resulting in a relatively low level of total error rate (Figure 5.4, Total Error Rate - SR - Trial Block 1). Thus, they tended to be more confident about the text they typed, and conducted relatively fewer gaze shift for proofreading (Figure 5.4, Gaze Shift - SR - Trial Block 1), which further led to less immediate error correction (Figure 5.4, Immediate Error Correction - SR - Trial Block 1).

As the users getting familiar with the new keyboard, they gradually developed their memory of the key locations and use it for guiding a more efficient key-searching (Figure 5.4, Gaze Path per Character - SR - Trial Block 2 and 3, decreasing gaze path per character) and more direct finger movement to the keys (Figure 5.4, Finger Efficiency - SR - Trial Block 2 and 3, increasing finer movement efficiency between keys). As a result, the typing speed increased (Figure 5.4, IKI - SR - Trial Block 2 and 3, dropping IKI reflects higher typing speed). Here we observed the speed-accuracy trade-off in typing: error rate increased with a higher typing speed (Figure 5.4, Total Error Rate - SR - Trial Block 2 and 3). In order to keep the typed text correct, users proofread more frequently (Figure 5.4, Gaze Shift - SR - Trial Block 2), leading to more immediate error correction (Figure 5.4, Immediate Error Correction - SR - Trial Block 2).

Although there is no guarantee that the users are able to reach the expert level in just 20 trials of sentence typing, we can still see a clear difference in their strategies as the users gaining experiences. In this final phase, users reached the highest level of typing speed during the experiment (Figure 5.4, IKI - SR - Trial Block 4, lowest IKI reflects the highest typing speed). Typing errors increased dramatically in this phase with the highest total error rate (Figure 5.4, Total Error Rate - SR - Trial Block 4). However, the benefit of fast typing speed not only compensated but also outweigh the costs of more errors and error corrections, leading to a generally faster typing speed on WPM (Figure 5.2).

5.5.2 Balance between Key-searching and Proofreading

Although users were forced to return to a novice state in both SR and DR, their behavior in those two conditions was different. Here we mainly focus on the discussion on the proofreading strategies. In SR, subjects were presented with a randomized keyboard which remained unchanged (static) throughout the trials of sentence typing. They were able to learn the key locations during use gradually. As with a Qwerty keyboard, to achieve a higher level of typing speed, users usually conduct operations in an overlapping manner, start searching for the next key even before the finger reaches the current target key [99]. However, in DR, all the letter keys on the keyboard were randomized after each keystroke. This means that searching for the next key before the current keystroke is useless as all the keys will be re-positioned. In such a situation, users could choose between either waiting for the keystroke to be finished and then search for the next key, or utilize this piece of time to proofread quickly. Our observations showed that users were more likely to shift their attention to the text input area for proofreading during this brief time interval, which further increased the possibility of finding an error immediately after it occurs. This explained why the immediate error correction (controlling the total error rate) was the highest under the condition of DR.

5.6 Summary

Users of today's technologies are constantly exposed to new or changed UIs. For facilitating ease of adoption and fluent adjustment to changes, it is important to understand how users obtain skills and change their behaviors during interaction with technologies. In this study, we manipulated typists' abilities between novice and skilled, observing the learning process as visible changes of eye and finger movements. At the core of our analysis of skill and its development is adaptation. We explain different amounts of skill—or learning of touchscreen typing—in terms of how users adapt to their mental resources, especially the internal memory they have of the UI. Via practice, this memory is updated and developed, permitting more efficient performance via behavioral adaptation that better exploits the evolved resources.

In some highlights of our results, the increase of finger movement efficiency and the decrease in the length of the path traveled by gaze are indicative of the aforementioned adaptation and an attempt to maximize gain. Designers can utilize this information to aid in an attempt to gauge their users' level of skill, as well as optimize interactions for these levels. For instance, knowing that a user has spent some time with a UI leads to expecting an improved finger movement efficiency, which can be accounted for in a more targeted design. Based on our description of changes in typing behavior due to gathering of experience, a question arises: what makes high-speed typing possible? Studies have shown that WPM values as high as 80 can be reached while typing with touchscreens [63]. The changes in visual-motor patterns described herein may shed light on what makes these fast typists perform. Hypothetically extrapolating from our results, the development of peak performance is a combination of strict motor control and visual guidance that adapts to it. As finger movements become more precise, less visual guidance is required for fast but relatively error-free pointing. This frees the vision for continuous checking of the typed message, allowing immediate correction of errors—unlike with those who type fast on physical keyboards [24]. Our results hint towards this explanation by showing, for skilled typing, improved finger efficiency and speed as well as increased proofreading time and less time spent with the gaze on the keyboard. This finding can be used to design a training regime for achieving faster typing, as we pinpoint that some of the fastest and most impactful ways to train are related to finger efficiency and gaze deployment.

The detailed analysis of the learning process reported here provides a basis of reference for HCI researchers and designers on how touch screen typists adapt their behavior and strategies, balancing between the need for efficient typing and correctness of the final input, according to the development of skill. This work encourages considering the user's typing skill and experience during UI design, and provides detailed information that can contribute to the development of better models of UI adaptation and skill.

CHAPTER 6

STUDY 4: INVESTIGATION ON THE STRATEGIES OF FAST TYPIST ON MOBILE DEVICES

6.1 Introduction

Mobile typing is usually considered as much slower than typing on a physical keyboard. During typing on the mobile device, frequent attention shift between the keyboard and text input area happens due to the lack of tactile feedback, which slows down the typing speed. Nevertheless, there still exist fast typists who can reach 100 WPM. Understanding their typing behavior and strategies may help the development of text input design. In this study, we recruited 4 participants with an average typing speed of 75.6 WPM, which is twice as fast as the normal-speed typist, and recorded their eye and finger movement during a transcription task. Preliminary results show that fast typists are more focused on the text input area than normal-speed typists. This brings the discussion of the shared strategies between physical keyboard typing and mobile typing.

6.2 Method

To capture the natural typing behavior of the fast typists, we designed an experiment in which participants use their own mobile devices and keyboard applications for transcription tasks. We use the native language of Finnish as the experiment was conducted at Aalto University, Finland. Typed sentences were representative of everyday messages. Participants typed with two thumbs with varying forms of typing aid or without it. An adaptable tracking system was designed to capture finger and eye movements as well as key-pressing data. To investigate whether the fast typists' typing skill was shared across different devices, we also collected typing data from the participants using a PC and physical keyboard.

6.2.1 Participants

We recruited 4 participants in total by an online typing test (4 females; age ranged from 19 to 21, with an average of 19.5). All participants reported to be right-handed and had normal or corrected vision (correction strength between -4 and +4). All the participants were native Finnish speakers. On average, they have been using touchscreen smartphones for 8.25 years (SD = 0.5) and spend 3 hours (SD = 1.0) typing on an everyday basis. All of the participants reported having been using two-thumb typing most of the time. One of the participants reported to have been using the typing aid of prediction list and have no experience in playing music with any instrument. The other three participants haven't been using any typing aid, but they had been practiced (or played) a musical instrument for at least two years.

The observed mobile typing speed ranged from 68.0 WPM to 79.3 WPM (M = 75.6, SD = 4.55), typing speed on the physical keyboard ranged from 71.2 WPM to 84.6 WPM (M = 78.1, SD = 4.9). Each participant was compensated with two movie tickets (total worth about 20 euros) for their time.

6.2.2 Experiment Design

Each subject typed 90 sentences, which were divided into three groups of 30 sentences under each of the conditions: 1) mobile keyboard with typing aids, 2) mobile keyboard without typing aids, and 3) physical keyboard. The typing aid includes auto-correction, smart punctuation, auto-capitalization, spelling check, word prediction, etc. During the experiment, the target sentences were firstly shown on the screen, then covered when the participants started typing to prevent them from distractions other than the typing task itself.

6.2.3 Material

We asked subjects to use their own smartphones for text input. A customized web-based typing application was developed for collecting key-pressing data and supporting the synchronization for the tracking system. Totally 120 sentences were selected from Yle News Archive Easy-to-read Finnish 2011-2018. The sentences were chosen to be short and simple enough for subjects to read and memorize before typing them. Specifically, we restricted the sentences by length and the frequency of words. All sentences contain at least one capital letter at the beginning and one symbol at the end. Numbers were not included in the sentences. The average sentence length was 31.2 characters (SD = 5.5).

6.2.4 Procedure

Firstly, participants were told that the purpose of the study was to analyze their eye and finger movements in mobile typing. During the experiment, they sat in a chair at an adjustable-height table with their own smartphones in their hands. Participants were free to adjust the height of the table to a comfortable level, and rest their hands and arms on it during the experiment. There were no physical constraints on the participants' hands or bodies. Participants were then instructed to practice and become familiar with the calibration procedure and typing task, with the eye-tracking glasses on their head. After a three-point calibration for the eye-tracking glasses, participants were asked to press four buttons on the screen, marked with the numbers 1 to 4 (Figure 6.1, left), in ascending order, for synchronization purposes.

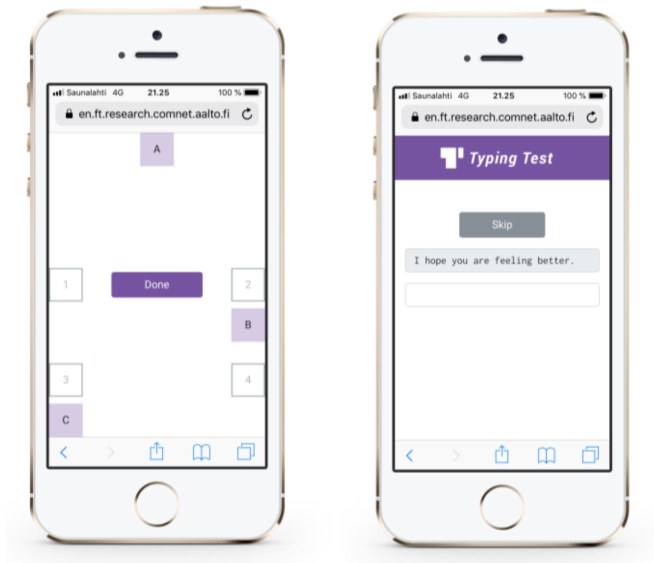


Figure 6.1 The interface of the experiment application: calibration screen (left), experiment screen with example sentences (right).

During the experiment, target sentences were shown one at a time on the screen, together with a text input area underneath the sentence. Participants had to read the sentence out loud, and then type It as fast and as accurately as possible, using two thumbs. The target sentence disappeared after the first key-press, so participants couldn't use visual copying but had to rely on their memory. In the “no typing aid” condition, error correction could be performed by either using the backspace button or by re-locating the cursor by touch before pressing backspace. In the “with typing aid” condition, auto-correction could be used. Participants finished typing a sentence by pressing the “Enter” key on the keyboard. In case they forgot the sentence while typing it, participants could click the “Skip” button and start the next sentence.

In each of the three conditions (with typing aid, without typing aid, and physical keyboard), we asked the participants to type two blocks of 15 sentences. After each block, we re-calibrated the eye tracker. At the beginning of each new condition, we had a practice block of 10 sentences. Three-minute rests were provided optionally in between each block of 15 sentences. After the experiment, we asked the subjects to fill a background questionnaire and recorded the sizes and models of their mobile devices.

6.2.5 Data Collection and Preprocessing

The data collected include eye movement, finger motion, and keypresses. For eye movements, we used SMI model 2W A eye-tracking glasses (60Hz at 3FPS). The glasses had infrared cameras tracking eye movements and a forward field camera to record the screen of the mobile device held in the hands. Participants with corrected vision had corresponding corrective lenses attached. The three-point calibration was done via the calibration screen, with the participant asked to focus one at a time on the rectangles marked with A, B, and C (Figure 6.1, left). Before the experiment, we attached a marker frame with four green markers to the smartphone for coordinate transformation purposes.

To track finger movement, we used an OptiTrack Prime 13 motion-capture system that provides 3D precision of up to 0.2mm at proximity. Two reflective markers were attached to the top-middle part of the nail of the two thumbs. Four reflective markers were also attached to the marker frame for tracking the smartphone position. The system was calibrated at the start of each 15 sentences, with the same calibration screen as for the eye-tracking device. The participants were asked to click on four buttons marked with numbers in order during the calibration process. The motion-tracking system labeled and tracked each marker during the experiment. In cases the tracker confuses the fingers with each other due to their proximity, we checked and corrected the data manually.

We extracted the coordinates from the raw data for the finger and eye movements and converted them into a common coordinate system. In the data, the upper-left corner of the screen is the origin (0,0,0), with x-axis values increasing toward the right of the device and y values from top to bottom. The distance from the screen facing upward is the positive z value. The unit in a datum refers to one millimeter.

6.3 Results

To keep the participants in a fast typing state, we asked them to use their own smartphone for the experiment, and track the movement of the phone by mounting a marker holder onto the phone. However, because the marker holder is not stable enough during the typing experiment, we got noisy data when trying to convert the gaze and finger coordinates into the same coordinate system. Thus, we dropped all the data of eye and finger movement, and try to understand the fast typists by analyzing the typing log.

6.3.1 Typing Speed

As the participants were invited based on an online typing test, their typing speed tends to be faster than normal. With the typing log during the experiment, we calculated the average typing speed for each participant. Mobile typing speed ranged from 68.0 WPM to 79.3 WPM (M = 75.6, SD = 4.55), typing speed on the physical keyboard ranged from 71.2 WPM to 84.6 WPM (M = 78.1, SD = 4.9). The distribution of typing speed per sentence (trial) presented as WPM was shown in Figure 6.2. Most of the sentences were typed at the speed of 80 WPM.

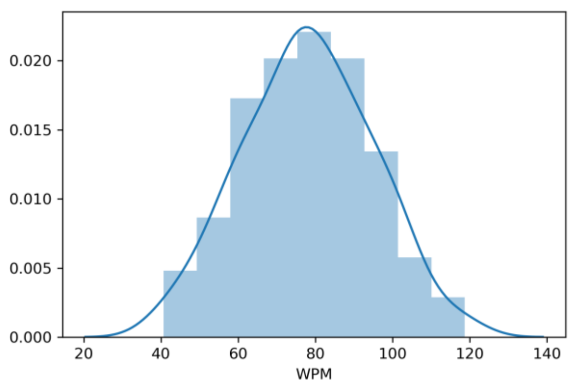


Figure 6.2 Distribution of the typing speed.

6.3.2 *Error and Error Correction*

Error correction in the experiment was mainly done by pressing the backspace key. Thus, when the error was overlooked for a while, the participant had to press the backspace key multiple times until the error and retype the correct letters. This operation is called “delayed error correction”. On the other hand, if the error was noticed and corrected with only one backspace immediately after the error occurred, the operation is called “immediate error correction”. Apart from those two types of error correction, we also looked at the number of backspaces pressed during typing on mobile devices.

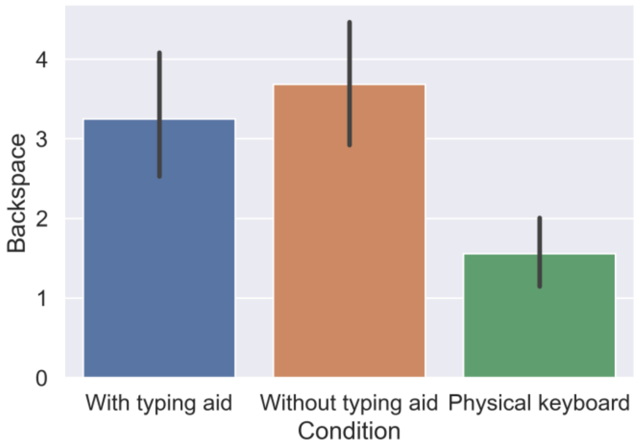


Figure 6.3 Number of backspaces pressed per sentence.

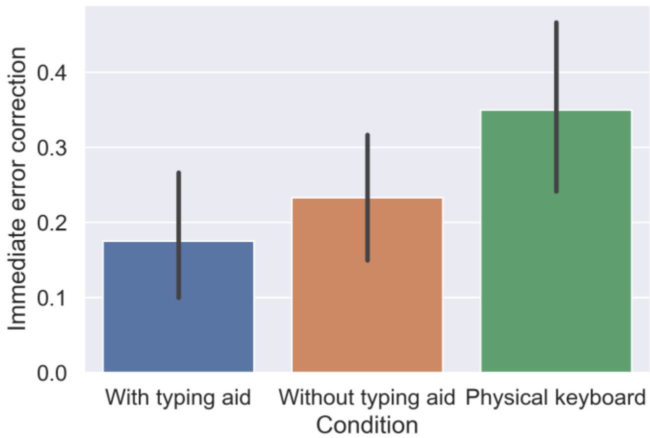


Figure 6.4 Number of immediate error corrections per sentence.

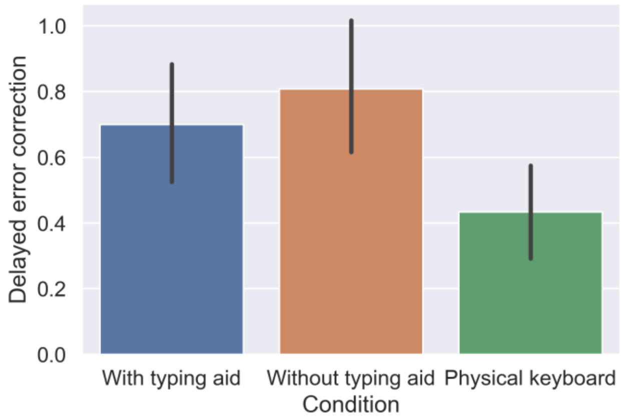


Figure 6.5 Number of delayed error corrections per sentence.

Typing without any typing aid from the keyboard or system led to more backspaces, immediate error correction, and delayed error correction, compared with using typing aids on the mobile device. Compared with typing on mobile devices, significantly fewer backspaces were pressed with a physical keyboard (Figure 6.3). Typing on mobile devices led to less immediate error correction (Figure 6.4) and more delayed error correction (Figure 6.5). To understand how error correction affects typing speed, we drew a scatter plot of the relationship between typing speed (WPM) and the average number of backspaces per sentence (Figure 6.6). As shown in Figure 6.6, when the typing speed increases, the participants pressed fewer backspaces, indicating a smoother input with fewer errors or error corrections.

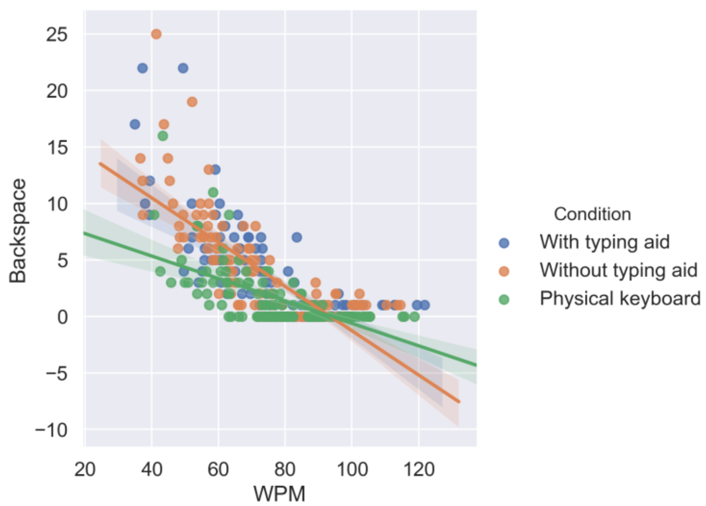


Figure 6.6 Relationship between typing speed and the number of backspaces.

6.3.3 *Attention during Typing*

We visualized the attention allocation of the participants during typing with heatmaps, and compared it with the interface of the experimental application. Here we illustrate some of the examples in Figure 6.7. Each heatmap indicates the distribution of visual attention for typing one sentence. As shown in Figure 6.7, attention during typing mainly focused on the text input area, indicating that the participants were monitoring the typed texts during typing. This result is different from our observation in study 2, when we analyzed the gaze data for normal-speed typists. We assume that the difference in attention allocation was related to fast typists' better finger control than normal-speed typists. There probably exists a robust mapping between the keys locations and the finger movements over the keyboard for the fast typists. And this mapping could help free their visual attention from guiding the finger movement to monitoring the typed text.

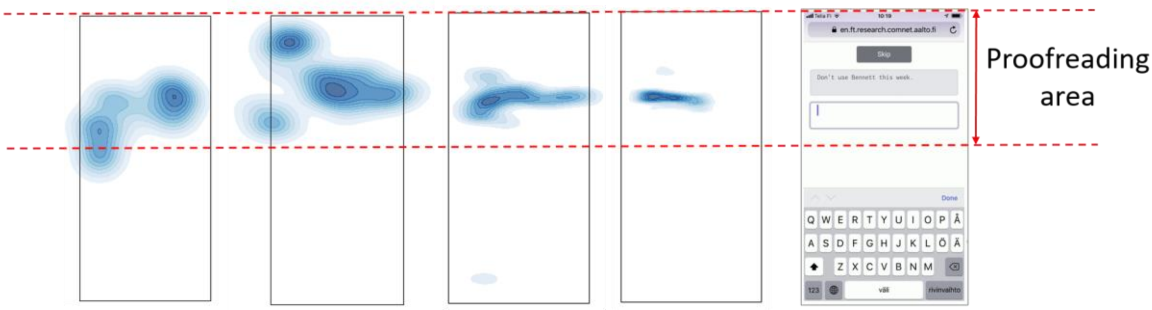


Figure 6.7 Heatmaps of visual attention locations during typing.

6.4 Summary

Typing at a relatively fast speed can benefit the keyboard users with higher communication efficiency. Many researchers are interested in improving the input techniques or underlying algorithms in order to lead the users to a higher typing performance. In this study, instead of changing the design of the keyboards, we focused on the existing fast typists who reached a relatively expert level of typing speed on a mobile device, and try to find the reason and operational patterns for the fast typing performance. Findings include: 1) the ability to type fast is shared between both mobile typing and physical keyboard typing, which indicates that the internal mechanism of typing is less affected by which devices or fingers used, as long as the users are familiar with the devices, 2) typing aids including auto-correction, smart punctuation, auto-capitalization, spelling check, word prediction can reduce the error happened and the number of error correction operations during typing on mobile devices, 3) compared with typing on mobile devices, typing on a physical keyboard leads to more sensitive detection of error and less delay on error corrections, 4) less error and error correction contributes to a higher typing speed. Although it may be hard for normal-speed users to reach the fast typing speed, understanding the difference between typing strategies of normal-speed and fast typists could still be inspiring. To summarize, to type fast, mobile device users could try to take advantage of the typing aids on smartphones, improve their finger movement accuracy, and be sensitive about the errors and correct errors as soon as it occurs.

CHAPTER 7

GENERAL CONCLUSIONS AND FUTURE RESEARCH DIRECTIONS

7.1 General Discussion and Contributions

Mobile devices support typing in a variety of scenarios. However, as one of the counter effects of mobility and convenience, being small and portable also brought difficulties in keeping a high typing efficiency. To be specific, as most of the typing was done using the Qwerty keyboard, mobile devices also embedded soft keyboards with the same layout. Putting at least 26 buttons on the lower part of the screen with the same relative locations made each of the keys smaller than the fingertips. While typing on such a keyboard, users have to deal with finger movement constraints and problems like occlusion and miss-touch. However, while there exist difficulties, mobile device users can still manage to reach a relatively acceptable typing speed with practice and growing expertise. As typing includes the control and cooperation of multiple systems, such as vision (i.e., eye movement) and motor (i.e., finger movement), we are interested to see how the eye and finger were controlled for achieving the typing task. Although literature extensively discussed the performance during typing, including typing speed, error rate, and so on, the detailed capture of movements for typing on a mobile device has not yet been done.

We aimed to capture, recognize, and summarize the eye-hand coordinative movement patterns during typing on mobile devices. This work revealed details of attention and movement control during typing, which provide references for future works on human behavior modeling on mobile devices. For example, we illustrated the attention shift between different areas on the interface, which indicated that attention is an undividable resource for sub-tasks in typing. Such findings, together with detailed data captured during the experiment, could contribute to the development of human behavior models.

7.2 Conclusions and Guidelines

As typing on mobile devices was derived from typing on a physical keyboard, we made a comparison with the previous work [24]. Compared with typing with a physical keyboard, in touchscreen typing, users spend more time looking at the keyboard area of the screen, and shift their attention more frequently between the text input area and the screen. Using two thumbs for typing tend to be faster than with only one index finger, as the finger movement can be done in an overlapped manner.

Remembering the keyboard layout benefits to a better typing performance. On the contrary, losing the memory of the key locations slows down the typing process. If a new layout was presented to a user, he/she needs to build internal mappings of the keys and locations during the typing process, in order to improve the efficiency of text input. In this process, the gaze will be mostly focused on the keyboard area as the keys need to be frequently searched for. Although typing is not fast in such a situation, users can keep a relatively lower error rate. Along with the findings of chapter 3, attention is also the key resource that user needs to allocate wisely, in order to reach a better performance.

Although faster typing speed can be achieved with practice and growing expertise for typing, it is still important to keep the error rate low. Generally, there are two strategies for keeping a lower error rate. First, if the key locations are well-memorized and the mappings between the finger movements and the keys are well-built, users can keep their attention mainly on the text input area. Keeping the attention on the typed text can help the users find the errors immediately after it was typed, and minimize the chance of any delayed error correction, which includes multiple key-pressings. Second, if the user is not familiarized with the keyboard, focusing eyes on the keyboard area will be a better option. By keeping attention on the keyboard area, users can quickly search for the target keys, and guide their fingers to them with relatively higher accuracy of key-pressing. By lowering the error rate, a higher typing speed can also be achieved. Thus, although there are strategies of how to allocate attention during typing, users still need to consider their expertise, and adjust their typing behavior accordingly.

7.3 Future Research Directions

Although we studied the user behavior of typing on mobile devices from different perspectives, including the effect of postures during typing (i.e., one-finger typing and two-finger typing), the learning of new keyboard layouts, and the features of fast typists, our understanding and findings were all based on the user behavior of typing on a Qwerty or randomized keyboard. There was no attempt to evaluate the typing behavior and its changes on an optimized keyboard or with other input methods like gesture typing, which was specifically designed for improving typing performance. Thus, as one of the future directions, typing behaviors on optimized keyboards with or without advanced input methods and error correction techniques could be captured and analyzed. Understanding eye and finger movements based on such new techniques could reveal not only the underlying reasons for the performance improvement but also potential directions for further improvements of the designs.

Another possible direction for future work is to build predictive models based on the detailed tracking data of eye and finger movements. Such models generate human-like data based on pre-defined constraints—leading to a more efficient evaluation of the newly designed typing interfaces without asking people to come to the lab. Furthermore, modeling of the eye and finger movement during typing can also contribute to adaptive user interfaces considering the limitations of both the users (e.g., users with motor control problems) and devices (e.g., multi-tasking with the same device leads to a smaller screen area for text input).

APPENDIX 1

BACKGROUND INFORMATION QUESTIONNAIRE FOR STUDY 1

Participant ID : ______ Name: __________ Date: _________ Time: ___________

1. Gender: A. Male B. Female

2. Age________

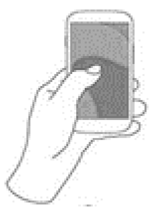
3. Which one is your dominant hand? A. Left hand B. Right hand

4. Which type of mobile phone are you using?
   A. Touchscreen smartphone, screen size is _____ inch
   B. Non-touchscreen smartphone
   C. Feature phone

5. How many years have you been using this type of mobile phone? _______years.

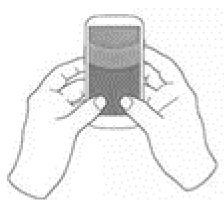
6. How long time do you usually use your phone per day?
   A. Less than 2 hours
   B. 2 hours to 4 hours
   C. 4 hours to 8 hours
   D. more than 8 hours

7. How do you usually use a smartphone? (multiple choice)

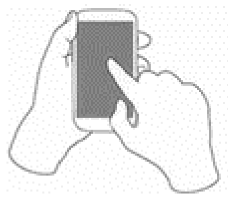


A. One-handed input with the thumb of the dominant hand

B. Input with the thumb of the dominant hand and hold the device with the non-dominant hand



C. Input with two thumbs



D. Input with the index finger of the dominant hand and hold the device with the non-dominant hand

8. What do you think of the difficulty of the pointing task in this experiment?

Very easy	1	2	3	4	5	Very hard

9. How much fatigue did you feel during the experiment?

Not tired at all	1	2	3	4	5	Very tired

10. How stressful you are during the experiment?

Not stressful at all	1	2	3	4	5	Very stressful

11. Please give any comments or feedback for the experiment.

APPENDIX 2

BACKGROUND INFORMATION QUESTIONNAIRE FOR STUDY 2 AND STUDY 3

Participant ID: _______________ *(filled by the experimenter)*
Age: ______ years

Circle one answer from each point.
Gender: male / female
Handedness: left / right

Complete the following statements by circling the most appropriate number. Choose only one answer from each point.

I use a computer (desktop or laptop) …

A. Several times a day
B. Once a day
C. A few times a week
D. A few times a month
E. Less than a few times a month
F. Not at all

I use touch - screen device(s) (e.g. mobile phone or tablet) ...

A. Several times a day
B. Once a day
C. A few a week
D. A few a month
E. Less than a few times a month
F. Not at all

Answer the following questions by writing an estimated number of hours spent on given activities.

How many hours per week do you type on a computer keyboard?
______ hours

How many hours per week do you type on a touch-screen device keyboard?
______ hours

APPENDIX 3

BACKGROUND INFORMATION QUESTIONNAIRE FOR STUDY 4

Participant ID: ______ Age: _____ Experiment Date: ________ Time: ________

1. Gender: male / female / other________

2. Hand-preference: left / right / both

3. Which kind of phone (model) are you using right now? __________________

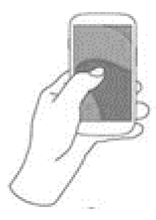
4. How many years have you been using a touchscreen smartphone? _______years.

5. What is your native language? __________________________

6. How often do you type in English (Finnish)?

   A. Several times a day
   B. Once a day
   C. A few times a week
   D. A few times a month
   E. Less than a few times a month
   F. Not at all

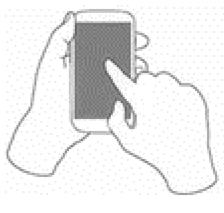
7. How do you usually type on a mobile phone? (multiple choice)



A. One-handed input with the thumb of the dominant hand

B. Input with the thumb of the dominant hand and hold the device with the non-dominant hand

C. Input with two thumbs



D. Input with the index finger of the dominant hand and hold the device with the non-dominant hand

8. Have you ever taken a typing course for desktop computers?

A. Yes B. No

9. How do you feel about your typing speed?

Very slow	1	2	3	4	5	6	7	Very fast

10. How many errors do you think you have during typing?

No errors	1	2	3	4	5	6	7	Many errors

11. Which typing aid do you use? (multiple choice)

A. Auto-error-correction
B. Word prediction
C. Swipe
D. No aids
E. Other_______________________

12. Which kind of keyboard do you use on the mobile phone?

A. System keyboard B. Keyboard app from the app store __________

13. Have you practiced or do you play any musical instrument?

A. No
B. Yes, I have practiced/played _________________for _______(years / months)

14. In what kind of situations you usually type with a mobile device? (multiple choice)

A. For work
B. For daily chats with friends
C. Others ____

APPENDIX 4

THE HOW-WE-TYPE-MOBILE DATASET

The HOW-WE-TYPE-MOBILE dataset was generated in study 2: Eye and Finger Movement Strategies in Mobile Typing. It contains eye-tracking, motion-tracking, and typing data of 30 participants typing regular Finnish sentences. It is publicly available at https://userinterfaces.aalto.fi/how-we-type-mobile/

The content of the dataset include:

1. Gaze data:
The x and y position of the gaze, recorded at 30 fps.
In the data, the upper-left corner of the screen is the origin (0,0), with x-axis values increasing toward the right of the device and y values from top to bottom.

2. Finger motion capture data:
The x, y, and z positions of the fingertips, recorded at 240 fps, frequency reduction to 60 fps during data processing. For one-finger typing, a marker is positioned on the index finger of the dominant hand. For two-finger typing, markers are positioned on the two thumbs.
In the data, the upper-left corner of the screen is the origin (0,0,0), with x-axis values increasing toward the right of the device and y values from top to bottom. The distance from the screen facing upward is the positive z value.

3. Typing log:
The x and y position of the touch, recorded at each typing event.
In the data, the upper-left corner of the screen is the origin (0,0), with x-axis values increasing toward the right of the device and y values from top to bottom.

4. Keyboard coordinates
The x-y coordinates of the center of each key on the soft keyboard, which had a Finnish Qwerty layout.

5. Background
Subjective responses to the survey, filled by the participants after the study.

6. Sentences
Sentence id and sentence.

More details about the study and its procedure can be found in the paper:

Xinhui Jiang, Yang Li, Jussi P.P. Jokinen, Viet Ba Hirvola, Antti Oulasvirta, Xiangshi Ren. How We Type: Eye and Finger Movement Strategies in Mobile Typing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, 2020.

BIBLIOGRAPHY

[1] A. Gunawardana, T. Paek, and C. Meek, “Usability guided key-target resizing for soft keyboards,” *Int. Conf. Intell. User Interfaces, Proc. IUI*, pp. 111–118, 2010.

[2] S. Zhu, T. Luo, X. Bi, and S. Zhai, “Typing on an invisible keyboard,” *Conf. Hum. Factors Comput. Syst. - Proc.*, vol. 2018-April, pp. 1–13, 2018.

[3] S. Azenkot and S. Zhai, “Touch behavior with different postures on soft smartphone keyboards,” *MobileHCI’12 - Proc. 14th Int. Conf. Hum. Comput. Interact. with Mob. Devices Serv.*, pp. 251–260, 2012.

[4] I. Scott MacKenzie and S. X. Zhang, “The design and evaluation of a high-performance soft keyboard,” *Conf. Hum. Factors Comput. Syst. - Proc.*, no. May, pp. 25–31, 1999.

[5] N. Henze, E. Rukzio, and S. Boll, “Observational and experimental investigation of typing behaviour using virtual keyboards on mobile devices,” *Conf. Hum. Factors Comput. Syst. - Proc.*, pp. 2659–2668, 2012.

[6] D. Buschek, J. Kinshofer, and F. Alt, “A Comparative Evaluation of Spatial Targeting Behaviour Patterns for Finger and Stylus Tapping on Mobile Touchscreen Devices,” *Proc. ACM Interactive, Mobile, Wearable Ubiquitous Technol.*, vol. 1, no. 4, pp. 1–21, 2018.

[7] J. H. Kim, L. S. Aulck, O. Thamsuwan, M. C. Bartha, C. A. Harper, and P. W. Johnson, “The effects of touch screen virtual keyboard key sizes on typing performance, typing biomechanics and muscle activity,” *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, vol. 8026 LNCS, no. PART 2, pp. 239–244, 2013.

[8] X. Yi, C. Yu, W. Shi, X. Bi, and Y. Shi, “Word clarity as a metric in sampling keyboard test sets,” *Conf. Hum. Factors Comput. Syst. - Proc.*, vol. 2017-May, pp. 4216–4228, 2017.

[9] P. O. Kristensson and K. Vertanen, “The inviscid text entry rate and its application as a grand goal for mobile text entry,” *MobileHCI 2014 - Proc. 16th ACM Int. Conf. Human-Computer Interact. with Mob. Devices Serv.*, pp. 335–338, 2014.

[10] P. D. Varcholik, J. J. Laviola, and C. E. Hughes, “Establishing a baseline for text entry for a multi-touch virtual keyboard,” *Int. J. Hum. Comput. Stud.*, vol. 70, no. 10, pp. 657–672, 2012.

[11] M. R. Zhang, S. Zhai, and J. O. Wobbrock, “Text entry throughput towards unifying speed and accuracy in a single performance metric,” *Conf. Hum. Factors Comput. Syst. - Proc.*, pp. 1–13, 2019.

[12] N. Banovic, V. Rao, A. Saravanan, A. K. Dey, and J. Mankoff, “Quantifying aversion to costly typing errors in expert mobile text entry,” *Conf. Hum. Factors Comput. Syst. - Proc.*, vol. 2017-May, pp. 4229–4241, 2017.

[13] N. Banovic, T. Sethapakdi, Y. Hari, A. K. Dey, and J. Mankoff, “The limits of expert text entry speed on mobile keyboards with autocorrect,” in *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI 2019*, 2019.

[14] P. Armstrong and B. Wilkinson, “Text entry of physical and virtual keyboards on tablets and the user perception,” *Proc. 28th Aust. Comput. Interact. Conf. OzCHI 2016*, pp. 401–405, 2016.

[15] H. Nicolau and J. Jorge, “Touch typing using thumbs: Understanding the effect of mobility and hand

posture," *Conf. Hum. Factors Comput. Syst. - Proc.*, pp. 2683–2686, 2012.

[16] V. Dhakal, A. M. Feit, P. O. Kristensson, and A. Oulasvirta, "Observations on typing from 136 million keystrokes," *Conf. Hum. Factors Comput. Syst. - Proc.*, vol. 2018-April, pp. 1–12, 2018.

[17] Hisao Yamada, "A historical study of typewriters and typing methods, from the position of planning Japanese parallel," *journal of Information Processing*. 1980.

[18] X. Bi and S. Zhai, "IJQwerty: What difference does one key change make? Gesture typing keyboard optimization bounded by one key position change from qwerty," *Conf. Hum. Factors Comput. Syst. - Proc.*, pp. 49–58, 2016.

[19] B. A. Smith, X. Bi, and S. Zhai, "Optimizing touchscreen keyboards for gesture typing," *Conf. Hum. Factors Comput. Syst. - Proc.*, vol. 2015-April, pp. 3365–3374, 2015.

[20] A. Oulasvirta *et al.*, "Improving two-thumb text entry on touchscreen devices," *Conf. Hum. Factors Comput. Syst. - Proc.*, pp. 2765–2774, 2013.

[21] M. D. Dunlop and J. Levine, "Multidimensional Pareto optimization of touchscreen keyboards for speed, familiarity and improved spell checking," in *Conference on Human Factors in Computing Systems - Proceedings*, 2012, vol. 1, pp. 2669–2678.

[22] X. Bi, B. A. Smith, and S. Zhai, "Quasi-Qwerty soft keyboard optimization," *Conf. Hum. Factors Comput. Syst. - Proc.*, vol. 1, pp. 283–286, 2010.

[23] M. D. Dunlop, N. Durga, S. Motaparti, P. Dona, and V. Medapuram, "QWERTH: An optimized semi-ambiguous keyboard design," *MobileHCI'12 - Companion Proc. 14th Int. Conf. Hum. Comput. Interact. with Mob. Devices Serv.*, pp. 23–28, 2012.

[24] A. M. Feit, D. Weir, and A. Oulasvirta, "How we type: Movement strategies and performance in everyday typing," *Conf. Hum. Factors Comput. Syst. - Proc.*, pp. 4262–4273, 2016.

[25] M. Rieger and V. K. E. Bart, "Typing style and the use of different sources of information during typing: An investigation using self-reports," *Front. Psychol.*, vol. 7, no. DEC, pp. 1–12, 2016.

[26] R. Johansson, Å. Wengelin, V. Johansson, and K. Holmqvist, "Looking at the keyboard or the monitor: Relationship with text production processes," *Read. Writ.*, vol. 23, no. 7, pp. 835–851, 2010.

[27] P. Weill-Tessier, J. Turner, and H. Gellersen, "How do you look at what you touch? A study of touch interaction and gaze correlation on tablets," in *Eye Tracking Research and Applications Symposium (ETRA)*, 2016, vol. 14, pp. 329–330.

[28] L. Ren, A. Z. Khan, G. Blohm, D. Y. P. Henriques, L. E. Sergio, and J. D. Crawford, "Proprioceptive guidance of saccades in eye-hand coordination," *J. Neurophysiol.*, vol. 96, no. 3, pp. 1464–1477, 2006.

[29] D. Baldauf, M. Wolf, and H. Deubel, "Deployment of visual attention before sequences of goal-directed hand movements," *Vision Res.*, vol. 46, no. 26, pp. 4355–4374, 2006.

[30] J. Huang, R. W. White, and G. Buscher, "User see, user point: Gaze and cursor alignment in Web search," *Conf. Hum. Factors Comput. Syst. - Proc.*, pp. 1341–1350, 2012.

[31] D. J. Liebling and S. T. Dumais, "Gaze and mouse coordination in everyday work," *UbiComp 2014 - Adjun. Proc. 2014 ACM Int. Jt. Conf. Pervasive Ubiquitous Comput.*, pp. 1141–1150, 2014.

[32] B. A. Smith, J. Ho, W. Ark, and S. Zhai, "Hand eye coordination patterns in target selection," *Proc. Eye Track. Res. Appl. Symp. 2000*, no. August, pp. 117–122, 2000.

[33] H. J. Bieg, L. L. Chuang, R. W. Fleming, H. Reiterer, and H. H. Bülthoff, "Eye and pointer

coordination in search and selection tasks," *Eye Track. Res. Appl. Symp.*, no. June 2014, pp. 89–92, 2010.

[34] P. M. Fitts, "The information capacity of the human motor system in controlling the amplitude of movement," *J. Exp. Psychol.*, vol. 47, no. 6, pp. 381–391, 1954.

[35] X. Bi, Y. Li, and S. Zhai, "FFitts law: Modeling finger touch with Fitts' law," *Conf. Hum. Factors Comput. Syst. - Proc.*, pp. 1363–1372, 2013.

[36] J. Bergstrom-Lehtovirta and A. Oulasvirta, "Modeling the functional area of the thumb on mobile touchscreen surfaces," in *Conference on Human Factors in Computing Systems - Proceedings*, 2014, pp. 1991–2000.

[37] P. Parhi, A. K. Karlson, and B. B. Bederson, "Target size study for one-handed thumb use on small touchscreen devices," *Proc. 8th Conf. Human-computer Interact. with Mob. devices Serv. - MobileHCI '06*, p. 203, 2006.

[38] M. E. Mott and J. O. Wobbrock, "Cluster touch improving touch accuracy on smartphones for people with motor and situational impairments," *Conf. Hum. Factors Comput. Syst. - Proc.*, pp. 1–14, 2019.

[39] E. Gowen and R. C. Miall, "Eye-hand interactions in tracing and drawing tasks," *Hum. Mov. Sci.*, vol. 25, no. 4–5, pp. 568–585, 2006.

[40] W. J. Horrey, C. D. Wickens, and K. P. Consalus, "Modeling drivers' visual attention allocation while interacting with in-vehicle technologies," *J. Exp. Psychol. Appl.*, vol. 12, no. 2, pp. 67–78, 2006.

[41] M. Oehl, C. Sutter, and M. Ziefle, "Considerations on efficient touch interfaces - How display size influences the performance in an applied pointing task," in *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, 2007, vol. 4557 LNCS, no. PART 1, pp. 136–143.

[42] K. S. Killourhy and R. A. Maxion, "Free vs. transcribed text for keystroke-dynamics evaluations," *ACM Int. Conf. Proceeding Ser.*, pp. 1–8, 2012.

[43] Y. Li, B. Zhang, Y. Cao, S. Zhao, Y. Gao, and J. Liu, "Study on the BeiHang keystroke dynamics database," *2011 Int. Jt. Conf. Biometrics, IJCB 2011*, pp. 1–5, 2011.

[44] A. Papoutsaki, A. Gokaslan, J. Tompkin, Y. He, and J. Huang, "The eye of the typer: A benchmark and analysis of gaze behavior during typing," *Eye Track. Res. Appl. Symp.*, pp. 1–9, 2018.

[45] T. Kujala, J. Mäkelä, I. Kotilainen, and T. Tokkonen, "The Attentional Demand of Automobile Driving Revisited: Occlusion Distance as a Function of Task-Relevant Event Density in Realistic Driving Scenarios," *Hum. Factors*, vol. 58, no. 1, pp. 163–180, Feb. 2016.

[46] J. A. Michon, "Critical View of Driver Behavior Models: What Do We Know, What Should We Do?," in *Human Behavior and Traffic Safety*, Boston, MA: Springer US, 1985, pp. 485–524.

[47] P. M. Greenfield, *Mind and media: The effects of television, video games, and computers*. Psychology Press, 2014.

[48] P. M. Greenfield, P. DeWinstanley, H. Kilpatrick, and D. Kaye, "Action video games and informal education: Effects on strategies for dividing visual attention," *J. Appl. Dev. Psychol.*, vol. 15, no. 1, pp. 105–123, Jan. 1994.

[49] D. H. Ballard, M. M. Hayhoe, F. Li, and S. D. Whitehead, "Hand-eye coordination during sequential tasks.," *Philos. Trans. R. Soc. Lond. B. Biol. Sci.*, vol. 337, no. 1281, pp. 331–339, Sep. 1992.

[50] U. Sailer, J. R. Flanagan, and R. S. Johansson, "Eye-hand coordination during learning of a novel

visuomotor task," *J. Neurosci.*, vol. 25, no. 39, pp. 8833–8842, Sep. 2005.

[51] Y. Guiard, H. B. Olafsdottir, and S. T. Perrault, "Fitts' law as an explicit time/error trade-off," *Conf. Hum. Factors Comput. Syst. - Proc.*, pp. 1619–1628, 2011.

[52] C. Holz and P. Baudisch, "Understanding touch," *Conf. Hum. Factors Comput. Syst. - Proc.*, pp. 2501–2510, 2011.

[53] S. Sarcar, J. P. P. Jokinen, A. Oulasvirta, Z. Wang, C. Silpasuwanchai, and X. Ren, "Ability-based optimization of touchscreen interactions," *IEEE Pervasive Comput.*, vol. 17, no. 1, pp. 15–26, 2018.

[54] S. Cao, A. Ho, and J. He, "Modeling and Predicting Mobile Phone Touchscreen Transcription Typing Using an Integrated Cognitive Architecture," *Int. J. Hum. Comput. Interact.*, vol. 34, no. 6, pp. 544–556, 2018.

[55] A. Sears, J. A. Jacko, J. Chu, and F. Moro, "The role of visual search in the design of effective soft keyboards," *Behav. Inf. Technol.*, vol. 20, no. 3, pp. 159–166, 2001.

[56] J. P. P. Jokinen, S. Sarcar, A. Oulasvirta, C. Silpasuwanchai, Z. Wang, and X. Ren, "Modelling learning of new keyboard layouts," *Conf. Hum. Factors Comput. Syst. - Proc.*, vol. 2017-May, pp. 4203–4215, 2017.

[57] P. Holleis, F. Otto, H. Hussmann, and A. Schmidt, "Keystroke-level model for advanced mobile phone interaction," *Conf. Hum. Factors Comput. Syst. - Proc.*, pp. 1505–1514, 2007.

[58] G. D. Logan and M. J. C. Crump, *Hierarchical Control of Cognitive Processes. The Case for Skilled Typewriting*, 1st ed., vol. 54. Elsevier Inc., 2011.

[59] T. A. Salthouse, "Perceptual, Cognitive, and Motoric Aspects of Transcription Typing," *Psychol. Bull.*, vol. 99, no. 3, pp. 303–319, 1986.

[60] M. D. Dunlop and M. M. Masters, "Pickup Usability Dominates: A Brief History of Mobile Text Entry Research and Adoption," *Int. J. Mob. Hum. Comput. Interact.*, vol. 1, no. 1, pp. 42–59, 2009.

[61] J. D. Gould, S. L. Greene, S. J. Boies, A. Meluson, and M. Rasamny, "Using a touchscreen for simple tasks," *Interact. Comput.*, vol. 2, no. 1, pp. 59–74, Apr. 1990.

[62] I. S. MacKenzie and R. W. Soukoreff, "Text entry for mobile computing: Models and methods, theory and practice," *Human--Computer Interact.*, vol. 17, no. 2–3, pp. 147–198, 2002.

[63] K. Palin, A. M. Feit, S. Kim, P. O. Kristensson, and A. Oulasvirta, "How do people type on mobile devices? Observations from a study with 37,000 volunteers," in *Proceedings of the 21st International Conference on Human-Computer Interaction with Mobile Devices and Services, MobileHCI 2019*, 2019.

[64] T. A. Salthouse, "Effects of age and skill in typing," *J. Exp. Psychol. Gen.*, vol. 113, no. 3, pp. 345–371, 1984.

[65] G. D. Logan, "Automatic control: How experts act without thinking," *Psychol. Rev.*, vol. 125, no. 4, pp. 453–485, 2018.

[66] M. Jeannerod, *The neural and behavioural organization of goal-directed movements*. Clarendon Press/Oxford University Press, 1988.

[67] E. Clarkson, K. Lyons, J. Clawson, and T. Starner, "Revisiting and Validating a Model of Two-thumb Text Entry," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2007, pp. 163–166.

[68] I. S. MacKenzie and R. W. Soukoreff, “A model of two-thumb text entry,” *Proc. - Graph. Interface*, pp. 117–124, 2002.

[69] P. O. Kristensson and A. Jameson, “Five challenges for intelligent text entry methods,” *AI Mag.*, vol. 30, no. 4, pp. 85–94, 2009.

[70] E. Hoggan, S. A. Brewster, and J. Johnston, “Investigating the effectiveness of tactile feedback for mobile touchscreens,” *Conf. Hum. Factors Comput. Syst. - Proc.*, pp. 1573–1582, 2008.

[71] J. Jokinen, “Touch Screen Text Entry as Cognitively Bounded Rationality,” 2017.

[72] P. Weill-Tessier and H. Gellersen, “Touch Input and Gaze Correlation on Tablets,” in *Intelligent Decision Technologies 2017*, 2018, pp. 287–296.

[73] K. Pfeuffer and Y. Li, “Analysis and modeling of grid performance on touchscreen mobile devices,” in *Conference on Human Factors in Computing Systems - Proceedings*, 2018, vol. 2018-April, pp. 1–12.

[74] Y. Lu, C. Yu, S. Fan, X. Bi, and Y. Shi, “Typing on split keyboards with peripheral vision,” *Conf. Hum. Factors Comput. Syst. - Proc.*, pp. 1–12, 2019.

[75] R. William Soukoreff and I. Scott Mackenzie, “Theoretical upper and lower bounds on typing speed using a stylus and a soft keyboard,” *Behav. Inf. Technol.*, vol. 14, no. 6, pp. 370–379, 1995.

[76] P. Quinn and S. Zhai, “A cost-benefit study of text entry suggestion interaction,” *Conf. Hum. Factors Comput. Syst. - Proc.*, pp. 83–88, 2016.

[77] S. J. Gershman, E. J. Horvitz, and J. B. Tenenbaum, “Computational rationality: A converging paradigm for intelligence in brains, minds, and machines,” *Science (80-.).*, vol. 349, no. 6245, pp. 273–278, 2015.

[78] R. L. Lewis, A. Howes, and S. Singh, “Computational rationality: Linking mechanism and behavior through bounded utility maximization,” *Top. Cogn. Sci.*, vol. 6, no. 2, pp. 279–311, 2014.

[79] jkielty, “Viewport, resolution, diagonal screen size and DPI for the most popular smartphones,” *DeviceAtlas*. 2019.

[80] K. Vertanen and P. O. Kristensson, “A versatile dataset for text entry evaluations based on genuine mobile emails,” *Mob. HCI 2011 - 13th Int. Conf. Human-Computer Interact. with Mob. Devices Serv.*, pp. 295–298, 2011.

[81] T. Cerni, M. Longcamp, and R. Job, “Two thumbs and one index: A comparison of manual coordination in touch-typing and mobile-typing,” *Acta Psychol. (Amst).*, vol. 167, pp. 16–23, 2016.

[82] J. O. Wobbrock, *Measures of Text Entry Performance*. Elsevier Inc., 2007.

[83] A. S. Arif, S. Kim, W. Stuerzlinger, G. Lee, and A. Mazalek, “Evaluation of a smart-restorable backspace technique to facilitate text entry error correction,” *Conf. Hum. Factors Comput. Syst. - Proc.*, pp. 5151–5162, 2016.

[84] J. R. Anderson, L. M. Reder, and C. Lebiere, “Working Memory: Activation Limitations on Retrieval,” *Cogn. Psychol.*, vol. 30, no. 3, pp. 221–256, 1996.

[85] M. J. Chapman, A. F. Healy, and J. A. Kole, “Memory load as a cognitive antidote to performance decrements in data entry,” *Memory*, vol. 24, no. 9, pp. 1182–1196, 2016.

[86] M. Yamaguchi and G. D. Logan, “Pushing typists back on the learning curve: Revealing chunking in skilled typewriting,” *J. Exp. Psychol. Hum. Percept. Perform.*, vol. 40, no. 2, pp. 592–612, 2014.

[87] C. M. Judd, J. Westfall, and D. A. Kenny, "Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem," *J. Pers. Soc. Psychol.*, vol. 103, no. 1, p. 54, 2012.

[88] U. Nikanne, "Finite Sentences In Finnish: Word Order, Morphology, And Information Structure," 2017.

[89] C. F. Jekel, G. Venter, M. P. Venter, N. Stander, and R. T. Haftka, "Similarity measures for identifying material parameters from hysteresis loops using inverse analysis," *Int. J. Mater. Form.*, vol. 12, no. 3, pp. 355–378, 2019.

[90] K. Witowski and N. Stander, "Parameter Identification of Hysteretic Models Using Partial Curve Mapping," in *12th {AIAA} Aviation Technology, Integration, and Operations ({ATIO}) Conference and 14th {AIAA}/{ISSMO} Multidisciplinary Analysis and Optimization Conference*, 2012.

[91] M. M. Botvinick, "Hierarchical reinforcement learning and decision making," *Curr. Opin. Neurobiol.*, vol. 22, no. 6, pp. 956–962, 2012.

[92] A. Oulasvirta, S. Tamminen, V. Roto, and J. Kuorelahti, "Interaction in 4-second Bursts: The Fragmented Nature of Attentional Resources in Mobile HCI," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 2005, pp. 919–928.

[93] S. Reyal, S. Zhai, and P. O. Kristensson, "Performance and user experience of touchscreen and gesture keyboards in a lab setting and in the wild," *Conf. Hum. Factors Comput. Syst. - Proc.*, vol. 2015-April, pp. 679–688, 2015.

[94] J. Bergstrom-Lehtovirta, A. Oulasvirta, and S. Brewster, "The effects of walking speed on target acquisition on a touchscreen interface," in *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services*, 2011, pp. 143–146.

[95] A. L. Smith and B. S. Chaparro, "Smartphone Text Input Method Performance, Usability, and Preference with Younger and Older Adults," *Hum. Factors*, vol. 57, no. 6, pp. 1015–1028, 2015.

[96] M. J. C. Crump and G. D. Logan, "Episodic contributions to sequential control: Learning from a typist's touch," *J. Exp. Psychol. Hum. Percept. Perform.*, vol. 36, no. 3, pp. 662–672, 2010.

[97] A. Howes, R. L. Lewis, and A. Vera, "Rational Adaptation Under Task and Processing Constraints: Implications for Testing Theories of Cognition and Action," *Psychol. Rev.*, vol. 116, no. 4, pp. 717–751, 2009.

[98] J. P. P. Jokinen, Z. Wang, S. Sarcar, A. Oulasvirta, and X. Ren, "Adaptive feature guidance: Modelling visual search with graphical layouts," *Int. J. Hum. Comput. Stud.*, vol. 136, no. February 2019, p. 102376, 2020.

[99] X. Jiang, Y. Li, J. P. P. Jokinen, V. B. Hirvola, A. Oulasvirta, and X. Ren, "How We Type : Eye and Finger Movement Strategies in Mobile Typing," *Proc. SIGCHI Conf. Hum. Factors Comput. Syst.*, no. January, pp. 1–14, 2020.

[100] J. R. Lewis, "Reaping the benefits of modern usability evaluation: The Simon story," in *Advances in Applied Ergonomics: Proceedings of the 1st International Conference on Applied Ergonomics--ICAE*, 1996, vol. 96, pp. 752–757.

[101] D. A. Norman and D. Fisher, "Why Alphabetic Keyboards Are Not Easy To Use: Keyboard Layout Doesn'T Much Matter.," *Hum. Factors*, vol. 24, no. 5, pp. 509–519, 1982.

[102] M. Hunter, S. Zhai, and B. A. Smith, "Physics-based graphical keyboard design," *Conf. Hum. Factors Comput. Syst. - Proc.*, no. April, pp. 157–158, 2000.

[103] A. Esteves, E. Velloso, A. Bulling, and H. Gellersen, "Orbits: Gaze interaction for smart watches using smooth pursuit eye movements," *UIST 2015 - Proc. 28th Annu. ACM Symp. User Interface Softw. Technol.*, no. 1, pp. 457–466, 2015.

[104] I. S. MacKenzie and S. X. Zhang, "An empirical investigation of the novice experience with soft keyboards," *Behav. Inf. Technol.*, vol. 20, no. 6, pp. 411–418, 2001.

[105] P. M. Fitts and M. I. Posner, "Human performance.," 1967.

[106] C. J. Chapman, "The learning curve in type writing," *J. Appl. Psychol.*, vol. 3, no. 3, pp. 252–268, 1919.

[107] N. Keith and K. A. Ericsson, "A Deliberate Practice Account of Typing Proficiency in Everyday Typists," *J. Exp. Psychol. Appl.*, vol. 13, no. 3, pp. 135–145, 2007.

[108] M. J. C. Crump and G. D. Logan, "Hierarchical Control and Skilled Typing: Evidence for Word-Level Control Over the Execution of Individual Keystrokes," *J. Exp. Psychol. Learn. Mem. Cogn.*, vol. 36, no. 6, pp. 1369–1380, 2010.

[109] M. Yamaguchi and G. D. Logan, "Pushing typists back on the learning curve: Memory chunking in the hierarchical control of skilled typewriting," *J. Exp. Psychol. Learn. Mem. Cogn.*, vol. 42, no. 12, pp. 1919–1936, 2016.

[110] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *arXiv Prepr. arXiv1406.5823*, 2014.

[111] A. Kuznetsova, P. B. Brockhoff, and R. H. B. Christensen, "lmerTest Package: Tests in Linear Mixed Effects Models ," *J. Stat. Softw.*, vol. 82, no. 13, 2017.

[112] R. W. Soukoreff and I. S. MacKenzie, "Metrics for text entry research: An evaluation of MSD and KSPC, and a new unified error metric," *Conf. Hum. Factors Comput. Syst. - Proc.*, no. 5, pp. 113–120, 2003.

ACKNOWLEDGMENTS

First, I would like to express sincere appreciation to Kochi University of Technology and Special Scholarship Program for giving me the opportunity to study and do research as a doctoral student. I would like to express my deep appreciation to my supervisor, Prof. Xiangshi Ren, who had created an excellent research environment in the Center for Human-Engaged Computing. I would not have learned so much and experienced so many research activities without his great support and guidance.

I would like to thank my committee members, Profs. Toru Kurihara, Yukinobu Hoshino, Yoshiaki Takata, and Shinichi Yoshida, for their time on my research and constructive feedback. I would like to sincerely thank Prof. Antti Oulasvirta and Dr. Jussi P.P. Jokinen for their tremendous support. I genuinely appreciate the effort they have put into the collaboration, providing constructive feedback with great synergy. The research methods and thinking deepened my understanding of how a good research project could be done. I am grateful to my friend and colleague Yang Li for his friendship, constructive comments, tremendous support, and countless research discussion and shared ideas. I would like to thank John Cahill for his help on English proofreading and his holistic comments. I feel lucky to meet Drs. Sayan Sarcar, Zhenxin Wang, William Delamare, and Kavous Salehzadeh Niksirat, as we not only had great discussions on research but also shared a good time in the Center for Human-Engaged Computing. Finally, I would like to thank our secretary at the Center for Human-Engaged Computing, Kyoko Hatakenaka, for her caring helps and happiness brought to us.

ACHIEVEMENTS

Publications

1. Jiang, X., Li, Y., Jokinen, J. P., Hirvola, V. B., Oulasvirta, A., & Ren, X. (2020, April). How We Type: Eye and Finger Movement Strategies in Mobile Typing. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1-14). (Orally presented at *CHI 2020 Japan Chapter* and *CHI Nordic 2020*)

2. Jiang, X., Jokinen, J. P., Oulasvirta, A., & Ren, X. (2021). Learning to Type with Mobile Keyboards: Findings with a Randomized Keyboard. *Computers in Human Behavior*. (Under review)

3. Jiang, X., Salehzadeh Niksirat, K., & Ren, X. (2021). Eye-tracking Correlates of Flow Experience in Video Games. *Computers in Human Behavior*. (To be submitted)

Grant

Travel grant (180,000JPY) from the Telecommunication Advancement Foundation for *ACM CHI Conference on Human Factors in Computing Systems (CHI 2020)*