FingerText: Exploring and Optimizing Performance for Wearable, Mobile and One-Handed Typing

DoYoung Lee
UNIST, Ulsan, Republic of Korea
dylee.issac@gmail.com

Jiwan Kim
UNIST, Ulsan, Republic of Korea
kjwan4435@gmail.com

Ian Oakley
UNIST, Ulsan, Republic of Korea
ian.r.oakley@gmail.com

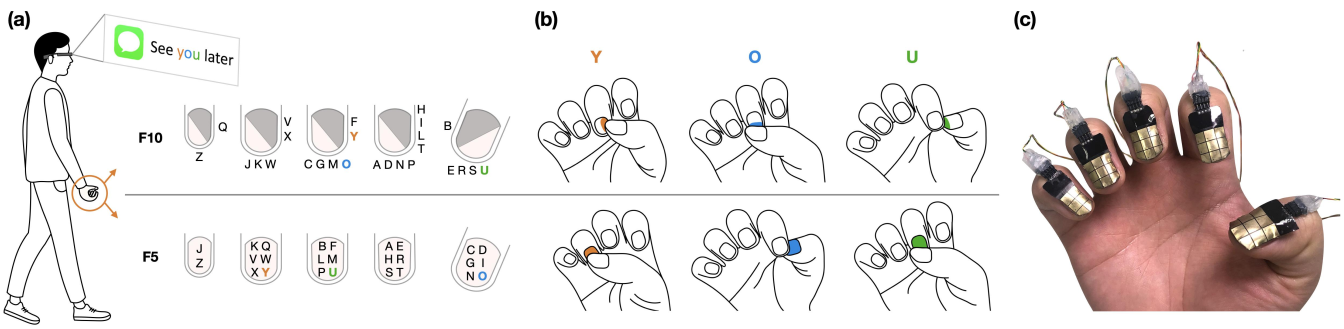

Figure 1: FingerText, a one-handed text entry system for touch sensitive nails. (a) shows two keyboard layouts: a ten key layout (*F10*) based on distinguishing between touches to the side and tip of each nail and a five key layout (*F5*) based on detecting a single touch event on each nail. (b) shows the sequence of inputs needed to type 'YOU' on both F10 and F5 layouts. (c) shows the set of nail touch sensors.

ABSTRACT

Typing on wearables while situationally impaired, such as while walking, is challenging. However, while HCI research on wearable typing is diverse, existing work focuses on stationary scenarios and fine-grained input that will likely perform poorly when users are on-the-go. To address this issue we explore single-handed wearable typing using inter-hand touches between the thumb and fingers, a modality we argue will be robust to the physical disturbances inherent to input while mobile. We first examine the impact of walking on performance of these touches, noting no significant differences in accuracy or speed, then feed our study data into a multi-objective optimization process in order to design keyboard layouts (for both five and ten keys) capable of supporting rapid, accurate, comfortable, and unambiguous typing. A final study tests these layouts against QWERTY baselines and reports performance improvements of up to 10.45% WPM and 39.44% WER when users type while walking.

CCS CONCEPTS

• **Human-centered computing** → **Keyboards**; **Text input**; **Mobile computing**.

KEYWORDS

wearable input, head-mounted display, optimization

ACM Reference Format:
DoYoung Lee, Jiwan Kim, and Ian Oakley. 2021. FingerText: Exploring and Optimizing Performance for Wearable, Mobile and One-Handed Typing. In *CHI Conference on Human Factors in Computing Systems (CHI '21), May 8–13, 2021, Yokohama, Japan.* ACM, New York, NY, USA, 15 pages. https://doi.org/10.1145/3411764.3445106

CHI '21, May 8–13, 2021, Yokohama, Japan

ACM ISBN 978-1-4503-8096-6/21/05...$15.00
https://doi.org/10.1145/3411764.3445106

1 INTRODUCTION

Text entry on wearable devices poses considerable challenges. Touch input spaces may be small [39], imprecise [55] or out of view [77]. Displays are also often small [67], or offset from input spaces [1]. Solutions to these problems involve techniques such as multi-stage character selection [34, 51], limited graphical feedback [46], bespoke gestural alphabets [77], and optimized keyboard layouts [38]. In addition, wearable devices inevitably target scenarios in which users are situationally impaired [69]: distracted [24], with one hand busy [20], or while mobile [57]. These situations demand wearable text entry systems that can be operated with one hand and while engaged in common activities such as walking. While researchers have begun to tackle situational impairments during wearable device use, such as enabling single-handed input [19], research on wearable device text entry while actually mobile is in its infancy, and remains focused on two-handed form factors such as touch input on smartwatches [1, 62].

However, mobility matters. Interaction on-the-go is a prevalent use scenario for smartphones [11] and the de facto design of such systems for stationary settings [30] has contributed not only to reduced input effectiveness and efficiency but also exacerbated social problems such as distracted walking [60] (or *twalking*), a potentially dangerous practice that has been banned in several US cities [10] due to perceptions of the risks it poses at pedestrian intersections. Explicitly designing for mobility aims to cater to, alleviate, or ameliorate these concerns—if we can better support and facilitate effective input, we can reduce the impact of user distraction. We argue these perspectives should also be applied to wearables. Indeed, while research on wearable interaction while on-the-go remains sparse, existing studies highlight unsurprisingly similar trends—mobility decreases input effectiveness [1, 13] and reduces performance in reading tasks [54], problems that can be mitigated, at least in part, through careful interaction design informed by data describing user performance in input tasks while mobile [57].

One promising design candidate for text entry in this space uses *intra-hand input*, defined as the combination of both thumb-to-finger [68] and finger-to-thumb touches. In such systems, a hand-worn or remote sensor, for example, a touch surface [31], depth sensor [59] or RFID tag [33] tracks contact between one or more fingers of one hand to register input events, most commonly taps and/or swipes. This type of in-hand input has been previously suggested as particularly appropriate for on-the-go interaction [80], an assertion supported by evidence indicating that performance of in-hand swipes (on a ring sensor) is robust to the disturbances caused by both walking and running [8]. The simplicity and ready availability of this modality have also led to diverse text-entry designs. Whitmire et al. [68], for example, distribute the characters of a full qwerty keyboard over the finger segments of both hands. A pair of touch-sensitive gloves tracks thumb touches to these regions and achieved a Words per Minute (WPM) of 16. Wong et al. [71] and Lee et al. [38] both propose broadly similar designs for single-handed use and report substantially lower WPMs: 5.42 and 6.47 WPM respectively. In a variation of this approach, Xu et al. propose a miniature keyboard spread over the distal phalanxes of either one [74] or both [73] index fingers; touch typing on these surfaces reached performance levels of between 13.3 WPM, for one hand, and an impressive 23.4 WPM for both. Finally, Fashimpaur et al. [15] use an external camera tracking system in conjunction with a Head-Mounted Display (HMD) to achieve WPMs of 12.54 for a two-handed keyboard design based on a probabilistic text entry system and touches to only the fingertips. While these text-entry projects have remained focused on stationary input, they effectively highlight the potential of thumb-to-finger input to support accurate, rapid, and fully wearable text entry.

We extend this work by considering the impact of, and designing expressly for, mobility. We first use the previously proposed form factor of a fingernail sensor system [31, 36] to assess input performance between sitting and walking conditions, logging both speed and accuracy (N=12). We contrast this data and our analysis shows no significant effects in performance in terms of speed or accuracy. We then combine these results with previously published data on the comfort of intra-hand finger touches [36] and simulations of the word-level accuracy of candidate keyboard layouts [40] in a quad-objective optimization process intended to generate key arrangements that can support high levels of user performance. Based on a balanced consideration of the results, we select two candidate keyboard layouts: *F5* and *F10*. *F5* is based on sensing single touches to each finger while *F10* assumes a higher fidelity system capable of distinguishing two touches per nail. Figure 1 (a) and (b) illustrate the layouts and how text input was performed in the system. We close by using our fingernail sensor system to evaluate both *F5* and *F10* layouts against QWERTY derived baselines in a word repetition task [5] while users are walking (N=16).

The contributions of this work are: 1) an evaluation of the impact of walking on intra-hand input performance; 2) an exploration of the design of keyboard layouts for intra-hand input that considers variations in sensing fidelity, in terms of the number of finger touches that can be detected, and uses computational methods to balance the competing concerns of speed, accuracy, comfort, and support for unambiguous word-level text input; 3) the *F5* and *F10* keyboard layouts selected to optimize performance and comfort for systems capable of detecting either one or two touches per finger and; 4) an evaluation of text entry performance using *F5* and *F10* while mobile, ultimately achieving WPMs, in a task simulating expert performance, of 31.3 and 25 respectively. This performance represents improvements of 9.47% and 10.45% over QWERTY baselines. The data, designs, and results we report, and methods we present, will help future researchers, designers and developers create more effective wearable text entry systems for mobile settings.

2 RELATED WORK

2.1 Wearable Text Entry

Text entry is a ubiquitous and challenging input task typically achieved, at high levels of performance, through large dedicated input devices such as keyboards. As computing devices have diversified into smaller mobile and wearable form factors, a considerable body of research has sought to enable rapid text entry on devices such as tablets [17], smart phones [58] and wearables such as finger sleeves [73] and rings [23]. The extremely small size and atypical input and output spaces of many wearables have led to a particularly wide range of proposals. A common approach involves adapting smartphone touch screen techniques to watches [51] or glasses [21]. Envisioning advanced finger trackers, researchers have also proposed external systems to track in-air finger strokes, thus supporting input actions that resemble those used in traditional keyboards and enabling rapid, accurate performance [15, 76]. Other proposals have sought to leverage the specific touch input capabilities of deployed wearable devices to create entirely novel schemes, such as a one-dimensional gesture input system [77]. More recently, text entry research has also begun to design for the situational impairments under which wearable devices are likely to be operated. WrisText [19], for example, targets one-handed use, recognizing that wearable device users may be engaged with everyday tasks such as holding a bag. Similarly, TipTex [74] and BiTipTex [73], based on miniature touch input surfaces mounted on the index finger(s), note that they may be accessible even when a user is holding a bag in the same hand as the device. Inspired by this work, we argue that successfully designing for mundane situational impairments, such as encumberment or walking, is likely to be a key factor in the eventual viability of any wearable text entry technique.

Text entry systems based on touches among the fingers of one hand are particularly relevant to this paper. A common approach has been to define touch input regions on the inner surfaces of the fingers. Whitmire et al. [68] applied this approach to both hands and a continuous input surface to create an unambiguous QWERTY keyboard. They demonstrated text entry speeds of 16 WPM after training, at a cost of encumbering both hands with a full glove input device. Other projects have explored a similar modality, but focused on single-handed use, arguing it is more practical in wearable settings. A major challenge in this work has been dealing with the reduced number of possible inputs this entails. Solutions include Jiang et al. [29]'s use of six input regions on the index and middle fingers with a two-stage input process (as in Zoomboard [51]) to uniquely specify characters, Lee et al. [38]'s use of nine input regions (i.e. adding the ring finger) each with three pressure-levels and [71]'s use of a similar 9 key layout with advanced word prediction techniques. While promising, these approaches have yielded somewhat limited text entry speeds of 5.42 [71], 6.47 [38] and 9.28 [29] WPM, after training. This is likely due to factors such as the time cost inherent in multi-stage selection processes (e.g., increased KeyStrokes Per Character (KSPC) [26]), reductions in input accuracy associated with eyes-free pressure input [61], the use of finger regions, such as the proximal phalanxes, that prior work has identified as uncomfortable [27], or reliance of QWERTY based layouts that may be poorly suited to the form factor of input on the fingers.

We build on this prior work by exploring typing on the nails. This finger region has not previously been studied in the context of text entry and we identify a number of reasons why it may be particularly suitable for this form of input. Firstly, the nails have been linked to improved comfort ratings [36], compared to those reported on the finger phalanxes. The nails may also offer greater input expressivity than the phalanxes, as they enable the thumb to be used as an additional touch surface [31], complementing input on the fingers. Each nail also supports several distinct touch regions [36] in close proximity, a fact that may support more rapid text entry times. Finally, nail wearables [63], unlike those mounted on the inner surfaces of the fingers, do not block tactile perception and intrinsically encumber the hand. In sum, nail input may be able to achieve comfortable, expressive, rapid, and unencumbered wearable text entry, a goal that is enticing and worthy of study.

2.2 Mobile Input

Mobility is a critical situational impairment for wearables. As with other mobile device form factors [30, 47], performance in tasks involving both viewing content [54] and performing input [13] on wearables drops while walking. However, work to understand and mitigate the impact of mobility on wearable interaction remains in its infancy. The majority of work to date [57], including in the area of text entry [1, 62], focuses on the relatively mature form factor of the smartwatch and deals with two-handed use — the watch is worn on one wrist and its screen tapped by the other hand. This input scenario closely models, and the reported results unsurprisingly follow, those for two-handed smartphone use [45]. While some work, such as Boldu et al. [8]'s touch-sensitive ring, which supports reliable input of swipes during a range of mobility conditions, highlights the potential for reliable wearable input while moving, we are not aware of work that empirically examines text entry while mobile on wearable devices other than the smartwatch. Indeed, researchers have recently identified exploring the performance of wearable typing systems while mobile as a key area for future work [29].

The studies in this paper are the first to address wearable, single-handed text entry while mobile. We seek to complement prior work on single-handed wearable text entry, which has focused on encumbered use — situations in which either one [19] or both hands [73] are occupied but the user is stationary. While this prior work demonstrates effective and elegant solutions for hands-busy use, we argue that many of the techniques it relies on, such as motion-based input [19] or micro-movements of the thumbs [73], will lose effectiveness in genuinely mobile settings. Walking will likely interfere with motion input, magnify the impact of encumbrance [48] and disturb acquisition of small targets [2], factors that suggest that the ability of these previously proposed techniques to support reliable and effective input while users are on-the-go may be extremely limited. On the other hand, we argue that the intra-hand touches we study, based on relatively large motions of all fingers relative solely to one another, may be highly resilient to the impact of walking [8]. A major goal of this paper is to establish the veracity of this claim.

2.3 Keyboard Optimization

In order to design a text entry technique, the letter assignment problem [16] refers to the process of allocating characters to input actions [78]. Key factors constraining this process are the number of characters that need to be supported and the number of input actions that are available. In wearable text entry systems, the number of input actions is typically less than the number of characters [27, 71, 74]. Input actions are therefore associated with multiple characters and word prediction techniques [14] are used to disambiguate input and enable accurate text entry. It is common to view the letter assignment problem as one of the multiple objectives, with different possible character arrangements resulting in different (and usually conflicting) performance profiles in terms of metrics such as text entry speed [43], input accuracy [19], comfort [38], word-level accuracy [40], or similarity to existing layouts [4], among others (see [16] for a full review). Multi-objective optimization, using processes such as evolutionary algorithms [58] or branch-and-bound integer programming [32], provides tools to help designers balance these concerns and select keyboard layouts that achieve a desired balance between objectives. They have been widely deployed to, for example, tweak QWERTY to boost performance without requiring retraining [5], reduce ambiguity in gesture keyboard designs [58] and improve the placement of infrequently used special characters [16]. In this paper, we leverage these methods by deploying an evolutionary algorithm to evaluate candidate key assignments on our wearable and mobile text input system in terms of input speed, comfort, accuracy, and the ability to uniquely specify words.

3 EXPERIMENTAL PLATFORM

All work in this paper used a set of five fingernails mounted capacitive sensors to track inter-hand touches, an approach also used by Lee et al. [36] and, in single nail systems, also by both Kao et al. [31] and Lee et al. [37]. We selected this approach as it supports relatively fine-grained tracking of up to five touches to each nail [31].

Furthermore, an informal comparison of published data characterizing the comfort of nail touches [36] against that for finger phalanx touches [27] suggests the former may be more comfortable for users. In addition, we expected that the reliability and robustness of capacitive sensing solutions would be greater than that of camera-based solutions—these remain at an early stage of development and while body-worn systems have been presented [59], even relatively recent work on text entry has relied on research-grade external motion capture systems [15] to support the high level of fidelity required for rapid, unambiguous sequences of input. In contrast the nail based capacitive sensing system we used is fully wearable and does not limit the user movement range.

We based our system on that presented by Lee et al. [36] and refer interested readers to this prior work for a more complete description. In brief, our sensor is composed of five separate modules, each attached to the nails of one hand. The modules consist of flexible PCBs (0.3 mm thick) mounted on commercial cosmetic nails and are attached to a user's nails using a standard fixative. Each nail has nine individual capacitive touch electrodes arranged in a square grid, bar the little finger whose narrower form supports six electrodes in two columns. Electrode sizes range from 3.8mm (on the fingers) to 4.8mm square (on the thumb). At the base of each nail sensor, each flexible PCB extends over the distal phalanx of the finger and contains an MPR121 micro-controller. This monitors the electrodes, reporting contact data in the form of an analog "touch heatmap" [72] at 100 Hz. During each sensor read, these heatmaps are processed to extract the largest contact region (via blob detection). We then calculate image moments to describe the region's centroid, angle, and dimensions. The centroid is considered the contact point on the nail. Each MPR121 is wired to a wrist-mounted Arduino MKR1010, via AWG32 gauge wires that do not restrict finger movement. The Arduino communicates with WiFi and UDP to a host computer. In terms of feedback, all studies reported in this paper used an Epson BT 200 Head Mounted Display (HMD) to present instructions and interfaces to users. This device provides a 23° field of view and a 30Hz update rate. All visual content on this device was controlled by the host computer via a wireless UDP connection. Figure 1 (c) shows a user's hand wearing the nail sensor.

4 METRICS STUDY

This study was designed to contrast performance with intra-hand touches between stationary and mobile conditions. We sought to document and characterize the performance changes that may occur with this form of input when users are mobile. In addition, it sought to complement existing data on the performance on individual intra-hand touches [27, 31, 36, 59] with data on how such touches are performed sequentially, as in a continuous process of typing. Data about the performance of all possible pairs of inputs in a system is required to support text entry optimization approaches that leverage bi-gram frequencies to create efficient and accurate input—they allow the fastest and most reliable input sequences to be assigned to the most common character pairings [43, 78]. To the best of our knowledge no prior work has captured a data set characterizing intra-hand input while users are mobile, nor one that systematically documents the performance of pairs of intra-hand finger touches. The study was approved by our university's IRB and was run in full compliance with governmental and institutional recommendations/restrictions for safety and social distancing.

4.1 Design

The study examined three independent variables: *pose* (sit/walk), *start-touch* and *end-touch*. Start-touch and end-touch had ten possible levels, each corresponding to the tip or side of one of the five fingernails. The study followed a fully repeated measures design: all participants completed trials in all conditions. Pose was balanced, with half the participants completing walking trials before sitting and the other half vice versa. Each unique combination of start- and end-touches formed a block within each pose condition. These blocks were randomly presented to the participants. Each block was composed of four repetitions of the start- and end-touch. The first repetition was discarded as practice. In addition, if a participant performed the requested pair of touches incorrectly, they were required to re-complete it. In total, this led to the retention of 600 correct trials per participant: two poses by ten start-touches by ten end-touches by three repetitions.

4.2 Participants

A total of 12 participants (seven male, five female, eleven right-handed and one mixed-handed) with a mean age of 23.08 (SD 1.67) completed this study using their dominant (right) hand. All were university students. On average, they self-rated themselves as highly familiar with computers (4.67/5.0, SD 0.65) and smartphones (4.83/5.0, SD 0.39) but only passingly familiar with virtual and augmented reality technology, such as the Epson smart glass system used in this work (2.0/5.0, SD 0.6). The study took approximately 50 minutes to complete and each participant was compensated with the equivalent of 30 USD in local currency.

4.3 Measures

We measured time and errors for pairs of touches. Time measurements were defined as the duration between initial contact with the start-touch finger region until initial contact with the end-touch finger region. Errors were defined as the number of times a participant failed to correctly select both start- and end-touch finger regions. In addition, during the walking conditions, we measured the overall distance participants travelled and used this to infer their average walking speed.

4.4 Procedure

The experiment took place in an unused class room. The study began with participants reading instructions, signing consent, and completing demographics. They then donned the study equipment, in the form of the nail sensor system on their right hand and the Epson HMD on their head. In order to provide control input, such as starting and stopping a trial, participants' also held a wireless mouse in their left hand, such that they could comfortably press its buttons. Next they performed all ten single nail touches to ensure they understand the input modality and that the system was worn comfortably. They then completed randomly presented (and never repeated) study trials until satisfied they were familiar with the format and instructions (five trials on average). The actual study trials then began. In the sit condition, participants sat in a chair

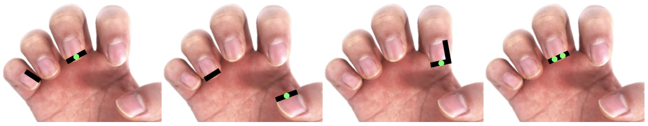

Figure 2: Example study instructions depicting sequential touches between different fingers/regions (left and center-left), different regions on the same finger (center-right) and, on the same region (right). A green dot indicates the start touch region and two green dots signify a double-tap.

without an arm rest, while in the walk condition they continuously walked a 30-meter figure-of-eight shaped route around a set of desks. They were requested to walk at a comfortable speed. A break of at least 5 minutes was enforced between the two pose conditions.

Each trial in the study followed a similar structure. First, participants clicked the wireless mouse and the experimental instructions were shown. These depicted the start- and end-touch regions, as highlights superimposed over a graphical image of a hand. Figure 2 shows example instructions. Participants then performed a pair of touches, following the shown instructions, and the trial ended and the next began. In between trials participants were able to rest if needed. Throughout the study, participants were asked to maintain a natural "arms down" posture, with their hand in free space near their waist or thigh. While we did not mandate eyes-free input, in practice this posture led, almost universally, to eyes-free performance of the intra-hand touches (see Figure 1, a).

4.5 Results and Discussion

We recorded a total of 7653 trials, including errors. We excluded 29 trials (0.38%) due to data loss caused by system failures, leaving a total of 7175 correctly completed trials and 449 error trials. To analyze time, we initially removed outliers by examining the correct trial data set as a whole, an intentionally conservative strategy, and excluded 150 trials (2.1%) with data over three standard deviations from the mean. We used mean imputation in the two (0.08%) cases when all a participant's data for a given condition was removed. We then plotted the data: Figure 3 (left) shows the main effects for time for all three dependent variables. Time data showed minor violations of normality in 8% of the individual conditions (i.e., 16 of the 200 combinations of pose, start- and end-touch). As ANOVA is widely viewed as robust to such distortions [18, 53], we analyzed time data using a single three-way repeated measures ANOVA, incorporating Greenhouse-Geisser sphericity corrections where indicated, on the variables of pose, start-touch, and end-touch. The significant results were a two-way interaction between start- and end-touch (F (81,891) = 7.55, p<0.001, $\hat{\eta}_G^2$=0.18) and individual main effects of both start- and end-touch (F (9,99) = 6.19, p<0.001, $\hat{\eta}_G^2$=0.029 and F (81,891) = 12.71, p<0.001, $\hat{\eta}_G^2$=0.10, respectively). The ANOVA did not find a significant effect of pose on the speed at which participants performed tasks.

Errors for the 200 individual conditions were not normally distributed — all individual conditions showed a median and interquartile range of zero. Furthermore, the data was predominantly discrete, with per participant error counts for individual conditions either zero (85.79% of trials), one (11.08%), or two (2.29%) and rarely greater (0.84%). This meant we were unable to apply Aligned Rank Transforms (ARTs) [70], a widely deployed technique to correct normality violations and deploy factorial parametric statistics, as they are designed for continuous data and inflate type I errors when applied to discrete data [41]. Accordingly, we opted to collapse the three individual variables and examine the main effects using non-parametric statistics. As this entails three separate tests, we applied an alpha value of 0.0167, equivalent to using Bonferroni correction. The collapsed data is plotted in Figure 3. Neither a Wilcoxon test on the pose variable ($W = 60.50$, $p = 0.52$), nor Freidman tests on start- and end-touch (respectively $\chi^2 = 17.6$, $p < 0.04$ and $\chi^2 = 10.8$, $p < 0.29$) led to significant differences. Finally, rather than leave the error data interactions entirely unexamined, and based on the presence of highly significant interactions in the time data, we collapsed pose and performed a two-way factorial RM ANOVA on start- and end-touch. The goal was to explore whether the interaction in the time data was also present in the error data. The results suggest it was: they revealed a significant interaction (F (81,891) = 1.92, p<0.001, $\hat{\eta}_G^2$=0.11) and a main effect of start-touch (F (9,99) = 2.06, p=0.04, $\hat{\eta}_G^2$=0.03) but not end-touch (F (9,99) = 0.33, p=0.96, $\hat{\eta}_G^2$=0.005). Given the normality violations in the data, these parametric results may have low validity. We include them as a speculative analysis due to the particular relevance of an interaction between these variables to work reported in this paper.

These results suggest that the main factor impacting performance was the relationship between the start- and end-touch finger regions—the interactions led to the largest effect sizes in the study. Rather than depend on which finger is touched, speed (and possibly accuracy) in the dual touch task we studied depended on the relationship between the start and end points. The confirms both our expectations and the general consensus in prior work [38, 43, 78]. We plot these relationships, for both time and errors, by reporting pairwise mean data in Figure 4. We opted not to conduct statistical pairwise comparisons on this data as both a large number of tests this would entail and the limited size/scope of our study would render these of questionable validity. Furthermore, the evidence in the interaction effects—that the performance of sequential intra-hand touches depends on the specifics of both start and end touches—is sufficient to support our main experimental objective and validates our goal of capturing data on the performance of sequential touches in order to support the design of optimized text entry input systems for intra-hand touch. Regardless, review of the raw pairwise means suggests several reassuringly expected trends. We observe that repeat selection of the same region (shown on the diagonal from left-top to bottom-right) leads to very strong performance; there are also noticeable benefits in sequential selection of two different regions on the same finger over two regions on different fingers (mean error rates of 2.58% versus 4.76% and task times of 520ms versus 583ms); the thumb and, particularly, the little finger, situated at the extremes of the hand, tend to yield lower performance and; there is a general (and anticipated [27, 36]) cluster of high performing regions on the index and middle fingers. These common-sense observations support the validity of the data we report.

In addition to these main results, we recorded an overall mean walking speed of 2.42 km/h (SD 0.52, min 1.41, and max 3.07), almost identical to that reported in prior accounts of mobile HMD use

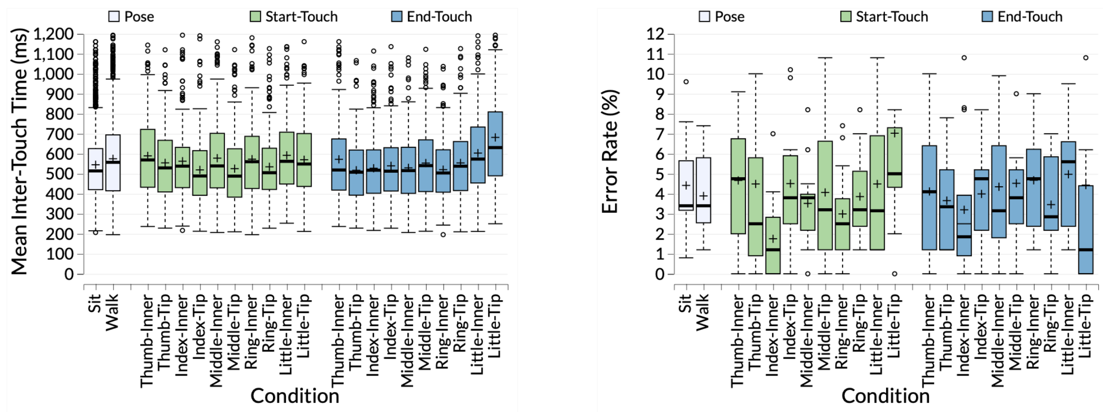

Figure 3: Box plots from the metrics study. Left shows time data from all conditions for all main effects. Right shows error data, collapsed to the individual independent variables, for all main effects. Means are marked by '+' symbols.

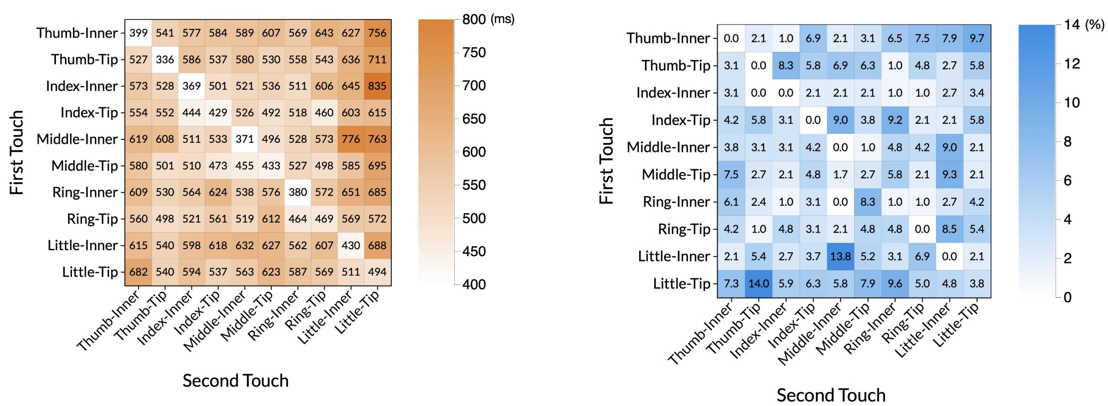

Figure 4: Raw mean results for all combinations of start- and end-touch in the metrics study. Left shows time data (ms) while right shows error data (%).

dealing with tasks such as reading [54]. While our study design does not support any formal comparisons with this literature, we suggest that the bespoke pictorial study instructions we used (see Figure 2) are unlikely to yield a lower mental load that of the highly practiced task of reading text. The fact that participants were able to both process our study instructions (shown on the HMD) and perform our input tasks while maintaining walking speeds at levels previously recorded during solely visual HMD use suggests that the intra-hand input we studied has limited impact on walking speed. While future studies would be required to confirm this preliminary suggestion, it does provide additional evidence of the viability of intra-hand input while on-the-go.

4.6 Conclusion

We draw two high level conclusions from the study. First and foremost, our analysis did not find significant effects of pose on performance. Numerically, both time and errors remained stable between sitting and walking poses. This, combined with the representative walking paces participants achieved, suggests that intra-hand input is a good candidate modality for wearable interaction while on-the-go. Users should be able to operate systems with intra-hand touches when walking with much the same ease as they can while seated. This is an extremely positive result given the widely documented performance reductions in other forms of touch input while walking [47]. We hope it spurs future work on this modality in mobile settings. Secondly, the relationship between start- and end-touches matters. This is of critical importance for the typing scenario we study as it suggests that keyboard layouts will need to take account of this relationship in order to support good user performance—simply applying existing (e.g., QWERTY) or default (e.g., alphabetical) layouts that do not consider this relationship will likely serve to limit the speed (and possibly the accuracy) with which skilled users are able to type. On the other hand, keyboards that ensure that commonly typed character sequences are achieved by a series of touches that can be performed rapidly and reliably may be able to boost performance to peak levels. The remainder of this paper explores how this can be achieved.

5 KEYBOARD LAYOUT OPTIMIZATION

Building on the results from the metrics study indicating the performance of sequential intra-hand input varies significantly based on the start- and end-touches, we conducted a keyboard layout optimization process to explore the range of possible designs. We considered hypothetical systems that are capable of detecting either single or a pair of touches to each nail and, to provide a rounded exploration of the space of possibilities in tractable computational time, used genetic algorithms, specifically NSGA-II [12], to achieve this. The goal was to generate five key (*F5*) and ten key (*F10*) layouts that are representative of optimal performance for each of these input scenarios. For 10 key layouts, we used the full set of data from our metrics study. In contrast, for five key layouts we used the subset of data from trials in the 25 conditions involving pairs of touches to nails tips. In both cases, our process was as follows: we first defined four metrics for assessing key layouts—speed; accuracy; comfort and; confusability. We selected these metrics to emphasize performance over concerns such as familiarity [38]. We then performed optimization processes for each metric individually in order to generate minimums and maximums for normalization. Next we ran multi-objective optimization using the normalized metrics. Finally, we used the resulting Pareto fronts, representing the sets of solutions in which no metric is dominant, as the source from which we selected final layouts for further study. We provide additional details on these processes in the sections below. Furthermore, Appendix A shows the mathematical formulations.

5.1 Metrics

5.1.1 Speed. Speed metrics for letter assignment problems typically combine temporal costs for arbitrary pairs of physical inputs with the bigram probabilities in a given text corpus. The goal is to design layouts that minimize the time taken to enter frequently occurring letter pairs. A common way to achieve this is via Fitts' law (e.g., as in the Fitts-Digraph model [43, 78]), an approach that models the time required to press pairs of keys in sequence as a function of the physical distance between them. However, for the type of intra-hand input we study, Fitts' law models are inappropriate. The physical distances between targets (i.e., fingertips) vary continuously due to diverse finger articulations such as correlated motions—involuntarily movements occur among other fingers during the intentional movement of one finger [76]. In addition, transitions that involve changes in the touching finger in addition to the touched finger (e.g., a touch to index with the thumb followed by a touch to thumb with the index) may result in higher costs than situations in which only the touched finger changes (e.g., the thumb touching the index then the middle fingers), irrespective of their proximity. As such, rather than use a Fitts' law derived model, we opted for the simple expedient of a function based on mean time costs for each possible pair of sequential inputs: the period between initial contact with the first finger region to initial contact with the second finger region. This approach is achievable due to the limited number of finger regions (5 or 10) considered in this work. Our speed function combines this data with bigram probabilities from Norvig [49] to model the overall time cost of a given letter assignment using the standard quadratic formulation of this problem [16].

5.1.2 Accuracy. We modelled accuracy using a similar mechanism to speed. We first assigned a cost for each possible pair of touches by multiplying their individual accuracy scores together. This represents a conservative view: if one touch in a pair is wrong, both are considered to be wrong. Using this strict metric during optimization was intended to ensure that more challenging input pairs were not assigned to high probability bigrams. We then combined pair input accuracy with bigram probabilities to model the overall accuracy cost of a given letter assignment: a second quadratic term.

5.1.3 Comfort. We used comfort ratings for nail touches from the literature [36] and followed Feit [16]'s recommendation to treat "ergonomic costs" such as comfort in terms of individual input events, rather than a property that emerges from pairs of events. The intuition here is that comfort ratings relate to the experience of specific input actions and cannot be meaningfully combined into an aggregate rating for a pair of actions. The cost function for comfort was therefore formulated as a linear term: it was simply based on the ratings for individual nail regions and the frequency with which individual characters occurred in our corpus.

5.1.4 Confusability. In key assignments in which multiple characters are assigned to each key, input is ambiguous. However, the sparsity of valid character sequences makes such systems effective: although each sequence of inputs can stipulate a range of possible character strings, only a small number correspond to actual words, meaning that word-level input remains relatively unambiguous [14, 19, 35, 38, 56, 75]. We included an assessment of the uniqueness of entered character sequences in our optimization process for a number of reasons. Firstly, the number of keys in our target *F5* layout is low—it requires a minimum of five to six characters assigned to each key. This will inevitably increase the number of valid words expressed by any given sequence of selections. In addition, our intuition was that accuracy, speed, and comfort metrics may result in grouping frequently selected letters, such as vowels, on the same high performing regions, an outcome that would likely substantially reduce the word-level accuracy of an ambiguous input system. Introducing a cost relating to the ability of layouts to accurately specify word-level input should be able to mitigate these problems and lead to layouts that balance the need to achieve a high level of input performance (e.g., that are fast, accurate, and comfortable) with the ability to unambiguously specify words.

A key challenge with this approach is substantial computational resources required to calculate the word-level accuracy of ambiguous character input [40]. While such computations are reasonable for a design process that evaluates the performance of tens to hundreds of manually selected key layouts [19], they are infeasible in the genetic algorithm driven optimization process we planned. To model this cost in a more tractable way we applied Lesher et al. [40]'s notion of pre-calculated confusability matrices. Based on a given text prediction algorithm, these matrices contain sums of the frequency with which all pairs of letters are mistakenly selected for each other in a given text corpus. During optimization, the cost of a particular key assignment is then calculated as the sum of matrix cells for all letters assigned to each key. We derived confusability matrices using Lesher et al. [40]'s optimal k-gram algorithm (implemented via word frequencies for the top thirty thousand most common words from Brants and Franz [9]) and a

322210 word corpus formed by combination of three mobile text entry data sets [50, 64, 65]. We used these matrices to assess the confusability of all layouts during our optimization process. This metric serves as a final (albeit relatively simple) quadratic term in our optimization process.

5.2 Normalization

In order to perform multi-objective optimization, all metrics need to be normalized so that variations in the units and scales each is expressed in do not unduly impact the process. A common way to achieve this is via approximating the minimum and maximum scores for each metric via independent optimization processes [58]. To achieve this, we used the NSGA-II algorithm [12] implemented in Pymoo [6] to determine the individual minimum and maximum scores for each of our four metrics for both *F5* and *F10* layouts. In total this involved 16 separate optimization processes: four metrics by two layouts by two endpoints. Each process involved 100 separate optimization runs, each configured with a population and offspring size of 300 and set to terminate after model improvements trailed off [7]—the default. We constrained the minimum number of letters that could be assigned to each key in both *F5* and *F10* to be one and the maximum to be, respectively, six and four. We derived these limits from the minimum numbers required to produce a valid arrangement (i.e. six keys in *F5*), and the four character/key limit used in prominent prior similar systems such as T9 [56].

5.3 Multi-objective Optimization

We used the same platform (NSGA-II/Pymoo) to conduct multiple objective optimization using the normalized metrics for both *F5* and *F10* layouts. We weighted all metrics equally and followed the same process used during normalization: 100 separate runs, population and offspring sizes of 300, default termination criteria and character assignments constraints of between one and six for *F5* and one and four for *F10*. We merged the results of all runs for both layouts, culling redundant solutions, to create four dimensional Pareto fronts for both *F5* and *F10* layouts. We visualize the fronts, composed of 3964 and 2297 layouts respectively, and in terms of their scores for all four de-normalized metrics in Figure 5. Inspection of these images reveals trade-offs between all metrics for the ten key case, but that the metrics of time, accuracy, and comfort tended to align for five keys. A clear trade-off was maintained between these three metrics and confusability throughout.

5.4 Results and Layout Selection

We selected optimal key layouts through a detailed review of those on the Pareto fronts. This was a multi-stage process. Firstly, to better interpret speed data, we calculated the projected WPM figures [42] for all Pareto front layouts. Secondly, to better contextualize the confusability metric we used, we augmented it by similarly calculating Gong et al. [19]'s *disambiguation score*. This metric expresses, for a given corpus of example words and dictionary of word frequencies, the mean rate at which a specific number of inputs on a specific key layout returns the intended example word with the highest probability. While calculating this metric is not computationally tractable during optimization, it is achievable for the relatively limited number of layouts on our Pareto fronts. We specifically calculated rates, assuming three entered keys, for which the correct word has the highest probability (*top1%*), is within the top three highest probabilities (*top3%*) and is within the top five highest probabilities (*top5%*). The goal was to better illustrate how a given layout may perform with respect to a text entry system capable of recommending lists of up to five high-frequency words for selection during typing. We note this analysis supported the validity of [40]'s confusability metric: over both *F5* and *F10* Pareto front layouts, Pearson correlations with top1, top3 and top5 scores showed very strong relationships of between 0.949 and 0.988. Finally, we created *Q5* and *Q10*, baseline qwerty-inspired designs for our five and ten key form factors (see Figure 6) and calculated scores for these on all metrics—see Table 1 for details. The goal of these activities was to provide a context for selecting novel designs.

We then reviewed all metrics for saliency. As shown in Figure 5, accuracy was generally high and did not vary substantially (95.6% to 99.4%) across either Pareto front; as such, we did not consider it during layout selection. In contrast, speed (450ms-563ms), comfort (3.73/5-4.67/5) and confusability (25512-353001) varied more considerably, with particularly clear trade-offs between layouts that achieve greater input speed and comfort and those that have reduced confusability. Based on the relatively strong performance of qwerty baselines in terms of confusability, we opted to select *F5* and *F10* layouts that emphasize improved user experience in terms of fast and comfortable input, while maintaining the best possible confusability (and thus top1, top3, and top5) scores. Figure 6 shows the final selected layouts and Table 1 their scores on all metrics. We note that both *F5* and *F10* target 10% or greater improvements in both speed and comfort over QWERTY designs. We also highlight that while the *F10* layout achieves relatively strong performance in terms of confusability, when compared to similar single-handed wearable input systems in the literature (e.g., 85.9% top1 and 95.3% top3 scores for Gong et al. [19]'s WrisText versus 88.1% and 97.7% for F10), the low number of keys on the *F5* inevitably compromises performance (to 70.1% and 88.6%) on this metric. In order to be usable, *F5* would require support from advanced word and sentence level prediction techniques. We selected it for further study in order to explore the extremes of user performance—to examine whether layouts that represent near peak values for comfort and speed actually provide the predicted benefits when real users actually type.

6 TEXT ENTRY STUDY

We conducted a final study to evaluate our *F5* and *F10* layouts against QWERTY baselines and validate how the metrics in the optimization process were reflected in the objective performance and subjective comfort of mobile wearable text entry. We used a word repetition task [5] as these can emulate expert levels of performance in relatively short study sessions and prior authors have shown they provide good estimates for performance in phrase level text entry tasks [3, 5, 66]. Due to the lack of differences observed in the metrics study, we opted not to re-examine the pose variable and all tasks in this study were conducted while participants were walking. The study was approved by our university's IRB and was run in full compliance with governmental and institutional recommendations/restrictions for safety and social distancing.

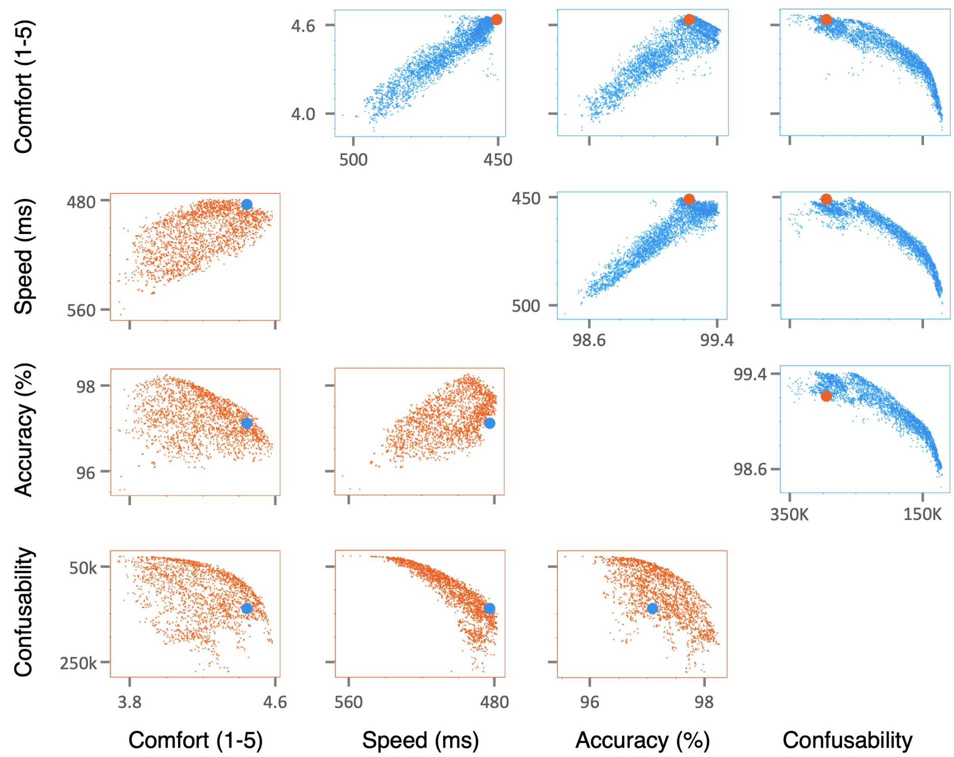

Figure 5: 2D projections of Pareto fronts from optimization processes for both five (in blue, top-right) and ten (in orange, bottom-left) key layouts for all six possible pairs of speed, accuracy, comfort, and confusability metrics. In each figure, top, right quadrants represent better performance. Large red dots indicate the locations of the F5 (in top-right charts) and F10 (in bottom-left chart) layouts. During layout selection, speed, and comfort metrics were emphasized.

Table 1: Performance metrics for selected F5 and F10 keyboards layouts and QWERTY derived baselines. *Norm* columns contain normalized scores, which are summed in the final column.

Keyboard	Speed (ms / WPM)			Accuracy (%)		Comfort (1-5)		Confusability (#)					Normalized
Layout	Value	WPM	Norm	Value	Norm	Value	Norm	Value	Norm	Top1%	Top3%	Top5%	Value Sum
QWERTY10 (Q10)	538	27.9	0.34	95.9	0.43	3.952	0.31	108863	0.25	91.3	98.9	99.7	1.332
F10	486	30.9	0.06	97.13	0.21	4.455	0.07	140154	0.34	88.1	97.7	99.1	0.685
QWERTY5 (Q5)	494	30.4	0.45	98.56	0.39	4.064	0.35	191122	0.17	81.7	95.1	97.8	1.369
F5	453	33.1	0.05	99.21	0.10	4.654	0.01	293959	0.42	70.1	88.6	93.7	0.582

6.1 Design

The study followed a repeated measures design with two independent variables: layout *type* (qwerty/optimized) and number *keys* (five/ten). All participants completed trials in all four layouts and both variables were balanced. Specifically, half the participants completed qwerty conditions before optimized and the other half vice versa. Within each of these groups, half of the participants always completed five key conditions before ten and the others ten before five. For each condition, participants completed 20 blocks, each containing seven trials, each involving typing one repetition of a single word. This design, and the word set, are taken from prior work [5, 79]. The word set is "the and you that is in of know not they get have were are bit quick fox jumps lazy on". It contains all the English letters and approximates both monogram and bigram frequencies. In total, we recorded 560 trials from each participant: four experimental conditions by 20 blocks by seven trials.

6.2 Participants

Sixteen participants (11 male, all right-handed, mean age of 24.5 (SD 3.1)) were recruited from the local university via social media channels. As in the metric study, they were highly familiar with computers (4.44/5.0, SD 0.73) and smartphones (4.75/5.0, SD 0.58) but only passingly familiar with virtual and augmented reality technology (2/5.0, SD 0.73). The study took approximately 90 minutes to complete and each was compensated with approximately 20 USD in local currency. Furthermore, to motivate participants to make input

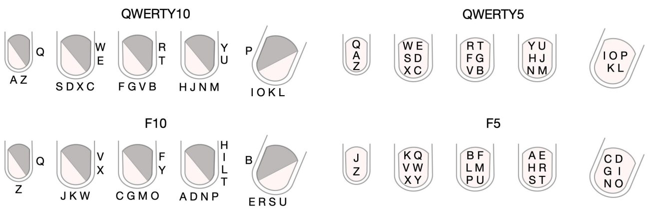

Figure 6: Keyboard layouts assuming either one or two touches each nail are possible. Figure illustrates QWERTY inspired baseline layouts and the *F5* and *F10* layouts produced via the multi-objective optimization process described in this paper. In all layouts, little finger is shown on the left and thumb on the right. In layouts based on two touches to each nail, characters are shown to the side or front of the nails to denote touches to the edge (darker area) or tip (lighter area) of the finger.

as quickly and accurately as possible, we awarded an additional 20 USD to two top performers.

6.3 Measures

We measured input speed, accuracy, and subjective comfort. For speed and accuracy, we used the standard metrics of Words Per Minute (WPM) [42] and Word Error Rate (WER)[5]. For comfort, participants completed a subjective assessment directly after every block in each condition. Specifically, participants provided a comfort rating on a one to five (uncomfortable to comfortable) scale, a process modelled on that used by Lee et al. [36], the source of the original comfort ratings used in this work. In addition, we once again logged the distance walked in each condition in order to subsequently calculate the walking speed.

6.4 Procedure

The study procedures, by and large, followed those used in the metrics study - participants walked around a figure of eight in the same classroom wearing the same equipment - nail sensor on their dominant (right) hand, HMD and wireless mouse. The study began by showing a visualization of the first layout (as illustrated in Fig 8) each participant was to experience and explaining how the characters were mapped to the finger regions. In particular, in the case of the Qwerty keyboards, additional explanations were given to ensure participants recognized the layout. Then they entered several (max five) randomly presented words to familiarize themselves with the system and then started the first condition. Each condition involved display of the 20 blocks in random order. The seven trials in each block involved display of the same word on the HMD, which participants were instructed to type. For typing feedback, we followed prior work [19, 71] by showing the first character on a key for the initial keystroke (e.g., C for F5 thumb) and the most likely word prefix for all further keystrokes. In the study, as we sought to verify whether our optimized layouts resulted in increased performance and/or comfort in a simplified task (as in [5]), we did not provide facilities for correcting errors. Rather, participants were simply instructed to type as rapidly and accurately as possible, without rectifying any mistakes. Individual trials were separated by breaks in which the participants were required to click the wireless mouse to move on. After completing a block, participants entered a comfort rating. As they were mobile, this frequent process was integrated into the wearable input system: participants tapped their thumb to indicate a very comfortable experience (5/5), their little finger to indicate a very uncomfortable one (1/5) and their other fingers to indicate the intermediary ratings. After completing each full condition, there was an enforced break (minimum 2 minutes).

6.5 Results

We first processed the 8960 trials recorded in the study by excluding the 330 outliers (3.3%) with a WPM over 3*IQR apart from the 1st and 3rd quartile. Forward fill imputation was used to ensure we retained complete pairwise data for all layouts and participants. We then plotted the WPM and WER data by repetition—see Figure 7 (left and center). These show substantial improvements during early repetitions, as the time spent in processes such as visual search reduces, and stable performance during latter repetitions as physical limitations become the constraining factors [5]. As our interest was in expert level performance, we conducted two-way RM-ANOVA on data from the final three repetitions for both WPM and WER metrics. These data are shown in Figure 7 (left-center and center-right). WPM showed no interaction, but significant main effects of *type* ($F(1,15)=4.59$, $p<0.05$, $\hat{\eta}^2_G=0.23$) and *keys* ($F(1,15)=79.1$ $p<0.001$, $\hat{\eta}^2_G=0.84$). WER showed an interaction ($F(1,15)=7.54$ $p<0.001$, $\hat{\eta}^2_G=0.33$) and also, again, main effects of *type* ($F(1,15)=7.56$, $p<0.05$, $\hat{\eta}^2_G=0.34$) and *keys* ($F(1,15)=65.2$ $p<0.001$, $\hat{\eta}^2_G=0.81$). In addition, we conducted RM-ANOVA on the mean comfort ratings (Fig 7, right). The interaction was significant ($F(1,15)=20.2$, $p<0.001$, $\hat{\eta}^2_G=0.02$), as was the main effect of *keys* ($F(1,15)=67.68$, $p<0.001$, $\hat{\eta}^2_G=0.14$).

Interpreting these results, we note the prevalent significant effects and particularly large effect sizes for *keys* indicate very robust performance improvements, at least in terms of speed and accuracy, when dropping to one key per finger from two: minimizing the number of keys per finger is a desirable approach for future systems. Furthermore, the moderate effect sizes for *type* suggest that the optimization process we conducted, and layouts we selected, were able to improve over baseline designs. Specifically, F5 improved by 9.47% (WPM) and 23.68% (WER) over Q5, and F10 by 10.45% (WPM) and 39.44% (WER) over Q10. These represent meaningful performance boosts. In addition, we saw few benefits of the QWERTY layouts in the early trials in each block—initial performance and learning curves were not noticeably better. We suggest participants were not able to map their knowledge of QWERTY to the form factor of their fingers and there may be few advantages to pursuing such layouts in the type of wearable system considered in this paper. The interaction effects can largely be explained by differences between *F10* and *Q10* that are weaker (WER), or absent (comfort), between *F5* and *Q5*. We conclude that improving layouts through optimization was somewhat less impactful in the five key case.

It is interesting the highlight how the study results either confirm or refute our expectations. Based on our optimization and layout selection process, we expected to see improvements in speed and comfort but not accuracy over qwerty baselines. While speed improvements materialized, these also translated into unexpected benefits in terms of accuracy, but comfort was not strongly boosted.

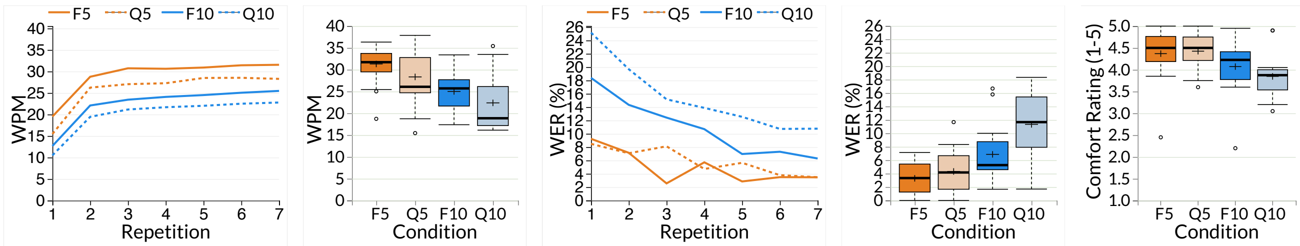

Figure 7: WPM (left-two), WER (center, center-right) and comfort (right) results from the text entry study. Line charts show how mean performance for each metric changes with repetition number. Box plots for WPM and WER are derived from data from the final three repetitions and represent expert performance. Comfort box plot summarizes all data.

Possible explanations for this may be that our metrics study failed to model the challenges, in terms of accuracy, of finger typing. Longer sequences of inputs may be required for this. In addition, the comfort ratings we used were extracted from a prior article dealing with individual touches to the nails [36]. These may not be directly applicable to the continuous input scenario we studied. As comfort is acknowledged to be an important factor in intra-hand touch input [29, 38], further work to understand how to best model comfort during wearable typing is currently required.

Finally, we note that the overall mean walking speed was 2.43 km/h (SD 0.61, min 1.49, and max 3.48), closely following that reported in the metrics study. This suggests that the word input task has a limited impact on walking speed and that our text entry system may be suitable for users on-the-go.

6.6 Discussion

It is worth contextualizing the results from this study with prior work that uses similar assessment methods [5], a similar single-handed text entry scenario [29, 38, 71] or addresses wearable text entry on-the-go [1, 62]. While no other work exists at the overlap of these three spaces, we are able to draw a number of interesting parallels and conclusions by examining each issue in turn. In terms of methods, Bi and Zhai [5] use a similar repeated word entry task for evaluating gesture typing keyboard layouts. Their findings align well with ours: text entry times stabilize from the third repetition and mean WER reach as high as 13.54%, figures similar to ours and due, at least in part, to the fact the task restricts participants from correcting errors. While it is not possible to draw strong performance parallels between such different input methods, these similar trends do suggest our methods did enable us to capture performance indicative of genuine expert use.

We can make more direct comparisons with work on in-hand text entry: using variations on single-handed taps to the finger phalanxes, Wong et al. [71], Lee et al. [38] and Jiang et al. [29] report input speeds of between 5.42 and 9.28 WPM. While the speeds reported in our study (between 22.38 and 31.3 WPM) clearly exceed these figures, this positive contrast must be considered in the context of study tasks used. These prior projects have sought to train users on keyboard layouts, over periods where that learning process is likely incomplete. As such, the performance they report does not represent the type of expert use we study—it is more likely "hunt and peck". In addition, our lab-based methods may overestimate the ability of skilled intra-finger typists. Regardless, our work does point towards appropriate strategies for maximizing performance in this area: reduce the number of different input actions by using ambiguous keyboards; focus on the most comfortable and easy to reach finger regions; optimize keyboard layouts; expect few benefits from using existing layouts such as QWERTY and; do not rely on prolonged or challenging input (e.g., pressure [38, 71] or multi-stage selection [29]).

Finally, in terms of mobility, prior work has established times for touchscreen typing on smartwatches: 9.019 WPM [1] in conjunction with smartglasses and 18-30 WPM using a range of input types on a standalone watch (handwriting recognition, keyboard and gesture keyboard) [62]. While this two-handed input differs greatly from that studied in this paper, we again note our one-handed WPMs contrast relatively well and suggest that in-hand typing interfaces may be a particular viable design candidate in this area. Rather than write on the tiny screen of your watch, it may be better to simply type with your hand [73, 74].

7 TEXT ENTRY SYSTEM DESIGN

Building on the positive results of this study, we developed a full text entry system for nail based in-hand input. This used the *F5* layout based on its strong performance in terms of speed, accuracy, and comfort. We added space and delete functionality using chords, or touches to a pair of fingers, and simple gestures [36]. Specifically, the space key was assigned to a simultaneous touch to index and middle fingernails and delete to touch to middle and ring nails. In order to facilitate rapid access, space and delete could also be redundantly accessed by swipes left and right over any nail. In addition, we developed a word suggestion and auto-complete system [19] capable of displaying the top five most likely words based on the currently typed characters. These suggestions were displayed above each finger in the HMD interface and could be selected by a dwelling (for 500ms) on the associated nail. Finally, alternative key layouts, such as for numbers or punctuation, were toggled by closing and re-opening the fist, an action that simultaneously triggers touches to all four fingernails. By integrating these diverse input modalities [36] (e.g., chord, dwell, and swipe) it was possible for the system to provide full keyboard functionality using only the nail sensors. Figure 8 shows an overview of this system. The next step for this project is to evaluate this full featured wearable text-entry system.

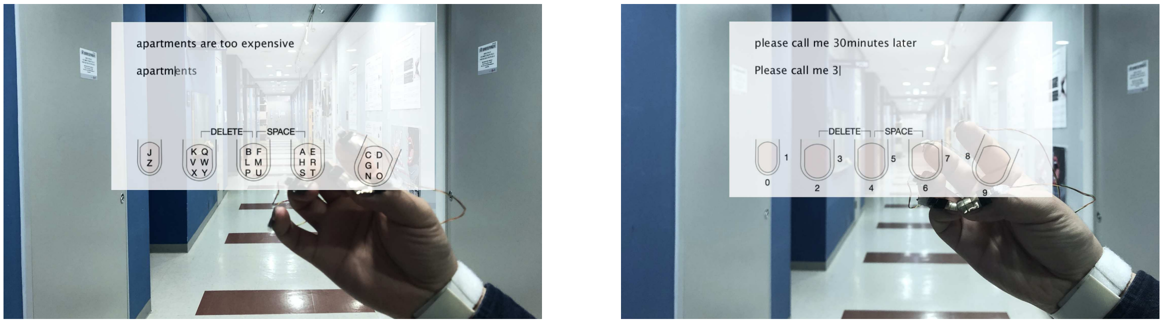

Figure 8: Nail touch based text entry system for mobile typing. It extends the system used in the study by integrating required keys (e.g., space, delete), a word selection interface and a hand close gesture to access a numeric keypad.

8 LIMITATIONS

A number of limitations impact our work. Many relate to its scope. At the highest level, we do not consider the broader implications of using wearable devices while mobile; we just focus on the typing experience. However, typing, or working with text in general, may exact a toll in terms of mental workload that would have implications in terms of, for example, safety. This may, in practice, preclude the design of such systems. While our work is motivated by the real world prevalence of mobile smartphone typing, the frequency with users genuinely would (or should) type on wearable devices is currently unknown: existing research on wearable device use patterns is relatively sparse and limited to the smartwatch form factor [28, 44]. Our studies could also be broader in scope. For example, we limited ourselves to two touches per finger, and focused on nail touches. While these are reasonable in terms of scope for a single paper, it would be interesting to explore systems that support more touches (e.g., five [31]), and/or other finger regions (e.g., the inner phalanxes [27]) and sensing systems [59]. Furthermore, our final study used only QWERTY derived baselines; another obvious baseline to study is alphabetical, a form used in prior predictive systems such as T9 [56]. Additionally, it would be valuable to compare performance directly against other current systems, such as head gaze or hand pointer-based keyboards. Longer multi-session text entry studies [22] capable of documenting learning curves and true expert performance with our layouts and system would also effectively complement the studies presented here. Furthermore, a study including the word suggestion and auto completion system and other keyboard functions (e.g., delete, space, and mode change) will generate valuable data about performance of more complex, realistic, and naturalistic typing tasks.

In addition, our optimization process could also be extended. For example, we could use alternative optimization approaches capable of guaranteeing the quality of solutions in terms of provably correct bounds [32]; we could apply weights, in a grid search pattern, to our multi-objective optimization process to more completely populate the Pareto set of solutions [58] and; we could conduct extended manual local searches in the regions around returned solutions—such structured local searches may improve the quality of the results [52]. Finally, we could also extend our treatment of mobility to consider other scenarios, such as travel or public transit [25], or conduct studies in-the-wild, in genuinely mobile settings (rather than the lab). Exploring the viability of wearable text entry in a broad range of mobile scenarios would be highly valuable.

9 CONCLUSION

This paper argues that text entry systems on wearables need to be designed to support mobility. As with smartphones, designing solely for stationary settings will result in systems that achieve low levels of performance when users inevitably opt to use them while mobile [30]. The resulting increased workload and frustration will lead to poor user experiences and potential societal harms—wearable device use while mobile may become unnecessarily hazardous. We explore the design of wearable text entry systems by documenting performance with the promising modality of intra-hand input. Our initial study confirms our expectations about the robustness of intra-hand input under mobile conditions. Contrasting walking and standing, our tests find no significant effects on performance for both input times and error rates. Building on these positive results, we conduct a multi-objective optimization process that seeks to balance the properties of input speed, accuracy, comfort, and ability to unambiguously specify words. We ultimately select two layouts: *F5* that relies on a single touch to each finger and *F10*, which assumes two touches to each finger can be detected. Our selection process emphasizes the speed and comfort of input while seeking to minimize reductions in the ability to unambiguously specify words. A second study shows these layouts provide performance improvements over QWERTY inspired baselines: they are up to 10.45% faster and 39.44% more accurate. While their ambiguity is increased (by 14.2% to 0.6% in terms of Gong et al. [20]'s disambiguation score) compared to QWERTY designs, we close by presenting the design of a word completion system that for our intra-hand input modality that we believe can mitigate these concerns and ensure users can achieve rapid, comfortable and accurate text entry performance using a wearable input device and while mobile.

ACKNOWLEDGMENTS

Thanks to all our study participants. This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT and Future Planning (2020R1F1A1070699).

REFERENCES

[1] Sunggeun Ahn, Seongkook Heo, and Geehyuk Lee. 2017. Typing on a Smartwatch for Smart Glasses. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces* (Brighton, United Kingdom) *(ISS '17)*. Association for Computing Machinery, New York, NY, USA, 201–209. https://doi.org/10.1145/3132272.3134136

[2] Joanna Bergstrom-Lehtovirta, Antti Oulasvirta, and Stephen Brewster. 2011. 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* (Stockholm, Sweden) *(MobileHCI '11)*. Association for Computing Machinery, New York, NY, USA, 143–146. https://doi.org/10.1145/2037373.2037396

[3] Xiaojun Bi, Ciprian Chelba, Tom Ouyang, Kurt Partridge, and Shumin Zhai. 2012. Bimanual Gesture Keyboard. In *Proceedings of the 25th Annual ACM Symposium on User Interface Software and Technology* (Cambridge, Massachusetts, USA) *(UIST '12)*. Association for Computing Machinery, New York, NY, USA, 137–146. https://doi.org/10.1145/2380116.2380136

[4] Xiaojun Bi, Barton A. Smith, and Shumin Zhai. 2010. Quasi-Qwerty Soft Keyboard Optimization. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '10)*. Association for Computing Machinery, New York, NY, USA, 283–286. https://doi.org/10.1145/1753326.1753367

[5] Xiaojun Bi and Shumin Zhai. 2016. IJQwerty: What Difference Does One Key Change Make? Gesture Typing Keyboard Optimization Bounded by One Key Position Change from Qwerty. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 49–58. https://doi.org/10.1145/2858036.2858421

[6] J. Blank and K. Deb. 2020. Pymoo: Multi-Objective Optimization in Python. *IEEE Access* 8 (2020), 89497–89509.

[7] J. Blank and K. Deb. 2020. A Running Performance Metric and Termination Criterion for Evaluating Evolutionary Multi- and Many-objective Optimization Algorithms. In *2020 IEEE Congress on Evolutionary Computation (CEC)*. 1–8.

[8] Roger Boldu, Alexandru Dancu, Denys J.C. Matthies, Pablo Gallego Cascón, Shanaka Ransir, and Suranga Nanayakkara. 2018. Thumb-In-Motion: Evaluating Thumb-to-Ring Microgestures for Athletic Activity. In *Proceedings of the Symposium on Spatial User Interaction* (Berlin, Germany) *(SUI '18)*. Association for Computing Machinery, New York, NY, USA, 150–157. https://doi.org/10.1145/3267782.3267796

[9] Thorsten Brants and Alex Franz. 2006. Web 1T 5-gram Version 1. *http://www.ldc. upenn. edu/Catalog/CatalogEntry. jsp? catalogId= LDC2006T13* (2006).

[10] Brian X Chen. 2019. Texting While Walking Is Dangerous. Here's How to Stop. *The New York Times* (Nov 2019). https://www.nytimes.com/2019/11/13/technology/personaltech/distracted-walking-twalking.html

[11] James Clawson, Thad Starner, Daniel Kohlsdorf, David P. Quigley, and Scott Gilliland. 2014. Texting While Walking: An Evaluation of Mini-Qwerty Text Input While on-the-Go. In *Proceedings of the 16th International Conference on Human-Computer Interaction with Mobile Devices & Services* (Toronto, ON, Canada) *(MobileHCI '14)*. Association for Computing Machinery, New York, NY, USA, 339–348. https://doi.org/10.1145/2628363.2628408

[12] K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan. 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* 6, 2 (2002), 182–197.

[13] David Dobbelstein, Gabriel Haas, and Enrico Rukzio. 2017. The Effects of Mobility, Encumbrance, and (Non-)Dominant Hand on Interaction with Smartwatches. In *Proceedings of the 2017 ACM International Symposium on Wearable Computers* (Maui, Hawaii) *(ISWC '17)*. Association for Computing Machinery, New York, NY, USA, 90–93. https://doi.org/10.1145/3123021.3123033

[14] Mark Dunlop and Andrew Crossan. 2000. Predictive Text Entry Methods for Mobile Phones. *Personal Technologies* 4 (08 2000). https://doi.org/10.1007/BF01324120

[15] Jacqui Fashimpaur, Kenrick Kin, and Matt Longest. 2020. PinchType: Text Entry for Virtual and Augmented Reality Using Comfortable Thumb to Fingertip Pinches. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI EA '20)*. Association for Computing Machinery, New York, NY, USA, 1–7. https://doi.org/10.1145/3334480.3382888

[16] Anna Maria Feit. 2018. *Assignment Problems for Optimizing Text Input.* Ph.D. Dissertation. Aalto University. http://urn.fi/URN:ISBN:978-952-60-8016-1

[17] Leah Findlater and Jacob Wobbrock. 2012. Personalized Input: Improving Ten-Finger Touchscreen Typing through Automatic Adaptation. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Austin, Texas, USA) *(CHI '12)*. Association for Computing Machinery, New York, NY, USA, 815–824. https://doi.org/10.1145/2207676.2208520

[18] Gene V Glass, Percy D. Peckham, and James R. Sanders. 1972. Consequences of Failure to Meet Assumptions Underlying the Fixed Effects Analyses of Variance and Covariance. *Review of Educational Research* 42, 3 (1972), 237–288. https://doi.org/10.3102/00346543042003237 arXiv:https://doi.org/10.3102/00346543042003237

[19] Jun Gong, Zheer Xu, Qifan Guo, Teddy Seyed, Xiang 'Anthony' Chen, Xiaojun Bi, and Xing-Dong Yang. 2018. WrisText: One-Handed Text Entry on Smartwatch Using Wrist Gestures. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–14. https://doi.org/10.1145/3173574.3173755

[20] Jun Gong, Xing-Dong Yang, and Pourang Irani. 2016. WristWhirl: One-Handed Continuous Smartwatch Input Using Wrist Gestures. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (Tokyo, Japan) *(UIST '16)*. Association for Computing Machinery, New York, NY, USA, 861–872. https://doi.org/10.1145/2984511.2984563

[21] Tovi Grossman, Xiang Anthony Chen, and George Fitzmaurice. 2015. Typing on Glasses: Adapting Text Entry to Smart Eyewear. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Copenhagen, Denmark) *(MobileHCI '15)*. Association for Computing Machinery, New York, NY, USA, 144–152. https://doi.org/10.1145/2785830.2785867

[22] Aakar Gupta and Ravin Balakrishnan. 2016. DualKey: Miniature Screen Text Entry via Finger Identification. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 59–70. https://doi.org/10.1145/2858036.2858052

[23] Aakar Gupta, Cheng Ji, Hui-Shyong Yeo, Aaron Quigley, and Daniel Vogel. 2019. RotoSwype: Word-Gesture Typing Using a Ring. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* (Glasgow, Scotland Uk) *(CHI '19)*. Association for Computing Machinery, New York, NY, USA, 1–12. https://doi.org/10.1145/3290605.3300244

[24] Jibo He, Jason S. McCarley, Kirsten Crager, Murtuza Jadliwala, Lesheng Hua, and Sheng Huang. 2018. Does wearable device bring distraction closer to drivers? Comparing smartphones and Google Glass. *Applied Ergonomics* 70 (2018), 156 – 166. https://doi.org/10.1016/j.apergo.2018.02.022

[25] Eve Hoggan, Stephen A. Brewster, and Jody Johnston. 2008. Investigating the Effectiveness of Tactile Feedback for Mobile Touchscreens. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Florence, Italy) *(CHI '08)*. Association for Computing Machinery, New York, NY, USA, 1573–1582. https://doi.org/10.1145/1357054.1357300

[26] Joggi Hong, Seongkook Heo, Poika Isokoski, and Geehyuk Lee. 2016. Comparison of Three QWERTY Keyboards for a Smartwatch. *Interacting with Computers* 28, 6 (10 2016), 811–825. https://doi.org/10.1093/iwc/iww003 arXiv:https://academic.oup.com/iwc/article-pdf/28/6/811/7920055/iww003.pdf

[27] Da-Yuan Huang, Liwei Chan, Shuo Yang, Fan Wang, Rong-Hao Liang, De-Nian Yang, Yi-Ping Hung, and Bing-Yu Chen. 2016. DigitSpace: Designing Thumb-to-Fingers Touch Interfaces for One-Handed and Eyes-Free Interactions. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 1526–1537. https://doi.org/10.1145/2858036.2858483

[28] Hayeon Jeong, Heepyung Kim, Rihun Kim, Uichin Lee, and Yong Jeong. 2017. Smartwatch Wearing Behavior Analysis: A Longitudinal Study. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 60 (Sept. 2017), 31 pages. https://doi.org/10.1145/3131892

[29] Haiyan Jiang, Dongdong Weng, Zhenliang Zhang, and Feng Chen. 2019. Hifinger: One-handed text entry technique for virtual environments based on touches between fingers. *Sensors* 19, 14 (2019), 3063.

[30] Shaun K. Kane, Jacob O. Wobbrock, and Ian E. Smith. 2008. Getting off the Treadmill: Evaluating Walking User Interfaces for Mobile Devices in Public Spaces *(MobileHCI '08)*. Association for Computing Machinery, New York, NY, USA, 109–118. https://doi.org/10.1145/1409240.1409253

[31] Hsin-Liu (Cindy) Kao, Artem Dementyev, Joseph A. Paradiso, and Chris Schmandt. 2015. NailO: Fingernails as an Input Surface. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 3015–3018. https://doi.org/10.1145/2702123.2702572

[32] Andreas Karrenbauer and Antti Oulasvirta. 2014. Improvements to Keyboard Optimization with Integer Programming. In *Proceedings of the 27th Annual ACM Symposium on User Interface Software and Technology* (Honolulu, Hawaii, USA) *(UIST '14)*. Association for Computing Machinery, New York, NY, USA, 621–626. https://doi.org/10.1145/2642918.2647382

[33] Keiko Katsuragawa, Ju Wang, Ziyang Shan, Ningshan Ouyang, Omid Abari, and Daniel Vogel. 2019. Tip-Tap: Battery-Free Discrete 2D Fingertip Input. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology* (New Orleans, LA, USA) *(UIST '19)*. Association for Computing Machinery, New York, NY, USA, 1045–1057. https://doi.org/10.1145/3332165.3347907

[34] Junhyeok Kim, William Delamare, and Pourang Irani. 2018. ThumbText: Text Entry for Wearable Devices Using a Miniature Ring. In *Proceedings of Graphics Interface 2018* (Toronto, Ontario) *(GI 2018)*. Canadian Human-Computer Communications Society / Société canadienne du dialogue humain-machine, 18 – 25. https://doi.org/10.20380/GI2018.04

[35] Andreas Komninos and Mark Dunlop. 2014. Using a smart-watch as a input device for text. *IEEE Pervasive Computing* 13 (01 2014).

[36] DoYoung Lee, SooHwan Lee, and Ian Oakley. 2020. Nailz: Sensing Hand Input with Touch Sensitive Nails. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376778

[37] DoYoung Lee, Youryang Lee, Yonghwan Shin, and Ian Oakley. 2018. Designing Socially Acceptable Hand-to-Face Input. In *Proceedings of the 31st Annual ACM Symposium on User Interface Software and Technology* (Berlin, Germany) *(UIST '18)*. Association for Computing Machinery, New York, NY, USA, 711–723. https://doi.org/10.1145/3242587.3242642

[38] Lik Hang Lee, Kit Yung Lam, Tong Li, Tristan Braud, Xiang Su, and Pan Hui. 2019. Quadmetric Optimized Thumb-to-Finger Interaction for Force Assisted One-Handed Text Entry on Mobile Headsets. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 3, 3, Article 94 (Sept. 2019), 27 pages. https://doi.org/10.1145/3351252

[39] Luis A. Leiva, Alireza Sahami, Alejandro Catala, Niels Henze, and Albrecht Schmidt. 2015. Text Entry on Tiny QWERTY Soft Keyboards. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 669–678. https://doi.org/10.1145/2702123.2702388

[40] Gregory Lesher, Bryan Moulton, and Jeffery Higginbotham. 1999. Optimal character arrangements for ambiguous keyboards. *IEEE transactions on rehabilitation engineering : a publication of the IEEE Engineering in Medicine and Biology Society* 6 (01 1999), 415–23. https://doi.org/10.1109/86.736156

[41] Haiko Luepsen. 2017. The aligned rank transform and discrete variables: A warning. *Communications in Statistics - Simulation and Computation* 46, 9 (2017), 6923–6936. https://doi.org/10.1080/03610918.2016.1217014 arXiv:https://doi.org/10.1080/03610918.2016.1217014

[42] I. Scott MacKenzie. [n.d.]. *2016. A note on calculating text entry speed. (2016).* https://www.yorku.ca/mack/RN-TextEntrySpeed.html

[43] I. Scott MacKenzie and Shawn X. Zhang. 1999. The Design and Evaluation of a High-Performance Soft Keyboard. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Pittsburgh, Pennsylvania, USA) *(CHI '99)*. Association for Computing Machinery, New York, NY, USA, 25–31. https://doi.org/10.1145/302979.302983

[44] Donald McMillan, Barry Brown, Airi Lampinen, Moira McGregor, Eve Hoggan, and Stefania Pizza. 2017. Situating Wearables: Smartwatch Use in Context. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 3582–3594. https://doi.org/10.1145/3025453.3025993

[45] Sachi Mizobuchi, Mark Chignell, and David Newton. 2005. Mobile Text Entry: Relationship between Walking Speed and Text Input Task Difficulty. In *Proceedings of the 7th International Conference on Human Computer Interaction with Mobile Devices & Services* (Salzburg, Austria) *(MobileHCI '05)*. Association for Computing Machinery, New York, NY, USA, 122–128. https://doi.org/10.1145/1085777.1085798

[46] Aske Mottelson, Christoffer Larsen, Mikkel Lyderik, Paul Strohmeier, and Jarrod Knibbe. 2016. Invisiboard: Maximizing Display and Input Space with a Full Screen Text Entry Method for Smartwatches. In *Proceedings of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Florence, Italy) *(MobileHCI '16)*. Association for Computing Machinery, New York, NY, USA, 53–59. https://doi.org/10.1145/2935334.2935360

[47] Josip Musić and Roderick Murray-Smith. 2016. Nomadic Input on Mobile Devices: The Influence of Touch Input Technique and Walking Speed on Performance and Offset Modeling. *Human–Computer Interaction* 31, 5 (2016), 420–471. https://doi.org/10.1080/07370024.2015.1071195 arXiv:https://doi.org/10.1080/07370024.2015.1071195

[48] Alexander Ng, John Williamson, and Stephen Brewster. 2015. The Effects of Encumbrance and Mobility on Touch-Based Gesture Interactions for Mobile Phones. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Copenhagen, Denmark) *(MobileHCI '15)*. Association for Computing Machinery, New York, NY, USA, 536–546. https://doi.org/10.1145/2785830.2785853

[49] Peter Norvig. 2009. Natural language corpus data. *Beautiful data* (2009), 219–242.

[50] D. R. O'Day and R. A. Calix. 2013. Text message corpus: Applying natural language processing to mobile device forensics. In *2013 IEEE International Conference on Multimedia and Expo Workshops (ICMEW)*. 1–6.

[51] Stephen Oney, Chris Harrison, Amy Ogan, and Jason Wiese. 2013. ZoomBoard: A Diminutive Qwerty Soft Keyboard Using Iterative Zooming for Ultra-Small Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) *(CHI '13)*. Association for Computing Machinery, New York, NY, USA, 2799–2802. https://doi.org/10.1145/2470654.2481387

[52] Antti Oulasvirta, Anna Reichel, Wenbin Li, Yan Zhang, Myroslav Bachynskyi, Keith Vertanen, and Per Ola Kristensson. 2013. Improving Two-Thumb Text Entry on Touchscreen Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) *(CHI '13)*. Association for Computing Machinery, New York, NY, USA, 2765–2774. https://doi.org/10.1145/2470654.2481383

[53] Roberto Refinetti. 1996. Demonstrating the Consequences of Violations of Assumptions in between-Subjects Analysis of Variance. *Teaching of Psychology* 23, 1 (1996), 51–54. https://doi.org/10.1207/s15328023top2301_14

[54] Rufat Rzayev, Paweł W. Woundefinedniak, Tilman Dingler, and Niels Henze. 2018. Reading on Smart Glasses: The Effect of Text Position, Presentation Type and Walking. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–9. https://doi.org/10.1145/3173574.3173619

[55] Tomoki Shibata, Daniel Afergan, Danielle Kong, Beste F. Yuksel, I. Scott MacKenzie, and Robert J.K. Jacob. 2016. DriftBoard: A Panning-Based Text Entry Technique for Ultra-Small Touchscreens. In *Proceedings of the 29th Annual Symposium on User Interface Software and Technology* (Tokyo, Japan) *(UIST '16)*. Association for Computing Machinery, New York, NY, USA, 575–582. https://doi.org/10.1145/2984511.2984591

[56] Miika Silfverberg, I. Scott MacKenzie, and Panu Korhonen. 2000. Predicting Text Entry Speed on Mobile Phones. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (The Hague, The Netherlands) *(CHI '00)*. Association for Computing Machinery, New York, NY, USA, 9–16. https://doi.org/10.1145/332040.332044

[57] Gaganpreet Singh, William Delamare, and Pourang Irani. 2018. D-SWIME: A Design Space for Smartwatch Interaction Techniques Supporting Mobility and Encumbrance. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (Montreal QC, Canada) *(CHI '18)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3173574.3174208

[58] Brian A. Smith, Xiaojun Bi, and Shumin Zhai. 2015. Optimizing Touchscreen Keyboards for Gesture Typing. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems* (Seoul, Republic of Korea) *(CHI '15)*. Association for Computing Machinery, New York, NY, USA, 3365–3374. https://doi.org/10.1145/2702123.2702357

[59] Mohamed Soliman, Franziska Mueller, Lena Hegemann, Joan Sol Roo, Christian Theobalt, and Jürgen Steimle. 2018. FingerInput: Capturing Expressive Single-Hand Thumb-to-Finger Microgestures. In *Proceedings of the 2018 ACM International Conference on Interactive Surfaces and Spaces* (Tokyo, Japan) *(ISS '18)*. Association for Computing Machinery, New York, NY, USA, 177–187. https://doi.org/10.1145/3279778.3279799

[60] Despina Stavrinos, Katherine W Byington, and David C Schwebel. 2011. Distracted walking: cell phones increase injury risk for college pedestrians. *Journal of safety research* 42, 2 (2011), 101–107.

[61] Craig Stewart, Michael Rohs, Sven Kratz, and Georg Essl. 2010. Characteristics of Pressure-Based Input for Mobile Devices. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Atlanta, Georgia, USA) *(CHI '10)*. Association for Computing Machinery, New York, NY, USA, 801–810. https://doi.org/10.1145/1753326.1753444

[62] Colton Turner, Barbara Chaparro, and Jibo He. 2020. Typing on a Smartwatch While Mobile: A Comparison of Input Methods. *Human Factors: The Journal of the Human Factors and Ergonomics Society* (02 2020), 001872081989129. https://doi.org/10.1177/0018720819891291

[63] Katia Vega and Hugo Fuks. 2014. Beauty Tech Nails: Interactive Technology at Your Fingertips. In *Proceedings of the 8th International Conference on Tangible, Embedded and Embodied Interaction* (Munich, Germany) *(TEI '14)*. Association for Computing Machinery, New York, NY, USA, 61–64. https://doi.org/10.1145/2540930.2540961

[64] Keith Vertanen and Per Kristensson. 2019. Mining, analyzing, and modeling text written on mobile devices. *Natural Language Engineering* (10 2019), 1–33. https://doi.org/10.1017/S1351324919000548

[65] Keith Vertanen and Per Ola Kristensson. 2011. A Versatile Dataset for Text Entry Evaluations Based on Genuine Mobile Emails. In *Proceedings of the 13th International Conference on Human Computer Interaction with Mobile Devices and Services* (Stockholm, Sweden) *(MobileHCI '11)*. Association for Computing Machinery, New York, NY, USA, 295–298. https://doi.org/10.1145/2037373.2037418

[66] Cheng-Yao Wang, Wei-Chen Chu, Po-Tsung Chiu, Min-Chieh Hsiu, Yih-Harn Chiang, and Mike Y. Chen. 2015. PalmType: Using Palms as Keyboards for Smart Glasses. In *Proceedings of the 17th International Conference on Human-Computer Interaction with Mobile Devices and Services* (Copenhagen, Denmark) *(MobileHCI '15)*. Association for Computing Machinery, New York, NY, USA, 153–160. https://doi.org/10.1145/2785830.2785886

[67] Dirk Wenig, Johannes Schöning, Alex Olwal, Mathias Oben, and Rainer Malaka. 2017. WatchThru: Expanding Smartwatch Displays with Mid-Air Visuals and Wrist-Worn Augmented Reality. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 716–721. https://doi.org/10.1145/3025453.3025852

[68] Eric Whitmire, Mohit Jain, Divye Jain, Greg Nelson, Ravi Karkar, Shwetak Patel, and Mayank Goel. 2017. DigiTouch: Reconfigurable Thumb-to-Finger Input and Text Entry on Head-Mounted Displays. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 3, Article 113 (Sept. 2017), 21 pages. https://doi.org/10.1145/3130978

[69] Jacob O. Wobbrock. 2019. *Situationally-Induced Impairments and Disabilities*. Springer London, London, 59–92. https://doi.org/10.1007/978-1-4471-7440-0_5

[70] Jacob O. Wobbrock, Leah Findlater, Darren Gergle, and James J. Higgins. 2011. The Aligned Rank Transform for Nonparametric Factorial Analyses Using Only Anova Procedures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vancouver, BC, Canada) *(CHI '11)*. Association for Computing Machinery, New York, NY, USA, 143–146. https://doi.org/10.1145/1978942.1978963

[71] Pui Chung Wong, Kening Zhu, and Hongbo Fu. 2018. *FingerT9: Leveraging Thumb-to-Finger Interaction for Same-Side-Hand Text Entry on Smartwatches*. Association for Computing Machinery, New York, NY, USA, 1–10. https://doi.org/10.1145/3173574.3173752

[72] Robert Xiao, Julia Schwarz, and Chris Harrison. 2015. Estimating 3D Finger Angle on Commodity Touchscreens. In *Proceedings of the 2015 International Conference on Interactive Tabletops & Surfaces* (Madeira, Portugal) *(ITS '15)*. Association for Computing Machinery, New York, NY, USA, 47–50. https://doi.org/10.1145/2817721.2817737

[73] Zheer Xu, Weihao Chen, Dongyang Zhao, Jiehui Luo, Te-Yen Wu, Jun Gong, Sicheng Yin, Jialun Zhai, and Xing-Dong Yang. 2020. BiTipText: Bimanual Eyes-Free Text Entry on a Fingertip Keyboard. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (Honolulu, HI, USA) *(CHI '20)*. Association for Computing Machinery, New York, NY, USA, 1–13. https://doi.org/10.1145/3313831.3376306

[74] Zheer Xu, Pui Chung Wong, Jun Gong, Te-Yen Wu, Aditya Shekhar Nittala, Xiaojun Bi, Jürgen Steimle, Hongbo Fu, Kening Zhu, and Xing-Dong Yang. 2019. TipText: Eyes-Free Text Entry on a Fingertip Keyboard. In *Proceedings of the 32nd Annual ACM Symposium on User Interface Software and Technology* (New Orleans, LA, USA) *(UIST '19)*. Association for Computing Machinery, New York, NY, USA, 883–899. https://doi.org/10.1145/3332165.3347865

[75] Xin Yi, Chun Yu, Weijie Xu, Xiaojun Bi, and Yuanchun Shi. 2017. COMPASS: Rotational Keyboard on Non-Touch Smartwatches. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (Denver, Colorado, USA) *(CHI '17)*. Association for Computing Machinery, New York, NY, USA, 705–715. https://doi.org/10.1145/3025453.3025454

[76] Xin Yi, Chun Yu, Mingrui Zhang, Sida Gao, Ke Sun, and Yuanchun Shi. 2015. ATK: Enabling Ten-Finger Freehand Typing in Air Based on 3D Hand Tracking Data. In *Proceedings of the 28th Annual ACM Symposium on User Interface Software & Technology* (Charlotte, NC, USA) *(UIST '15)*. Association for Computing Machinery, New York, NY, USA, 539–548. https://doi.org/10.1145/2807442.2807504

[77] Chun Yu, Ke Sun, Mingyuan Zhong, Xincheng Li, Peijun Zhao, and Yuanchun Shi. 2016. One-Dimensional Handwriting: Inputting Letters and Words on Smart Glasses. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '16)*. Association for Computing Machinery, New York, NY, USA, 71–82. https://doi.org/10.1145/2858036.2858542

[78] Shumin Zhai, Michael Hunter, and Barton A. Smith. 2002. Performance Optimization of Virtual Keyboards. *Human–Computer Interaction* 17, 2-3 (2002), 229–269. https://doi.org/10.1080/07370024.2002.9667315 arXiv:https://www.tandfonline.com/doi/pdf/10.1080/07370024.2002.9667315

[79] Shumin Zhai and Per Ola Kristensson. 2008. Interlaced QWERTY: Accommodating Ease of Visual Search and Input Flexibility in Shape Writing. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Florence, Italy) *(CHI '08)*. Association for Computing Machinery, New York, NY, USA, 593–596. https://doi.org/10.1145/1357054.1357149

[80] Shengdong Zhao, Pierre Dragicevic, Mark Chignell, Ravin Balakrishnan, and Patrick Baudisch. 2007. Earpod: Eyes-Free Menu Selection Using Touch Input and Reactive Audio Feedback. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (San Jose, California, USA) *(CHI '07)*. Association for Computing Machinery, New York, NY, USA, 1395–1404. https://doi.org/10.1145/1240624.1240836

A OPTIMIZATION PROCESS

This appendix includes objective functions used during the overall process and for each of the individual metrics. We recommend Feit [16] for a more comprehensive review and discussion of each of the individual metrics.

A.0.1 Objective function. The objective function for the overall optimization process is:

$$O(l) = (1 - time(l)) + acc(l) + comf(l) + (1 - conf(l)) \quad (1)$$

where our overall goal is to minimize the time costs, maximize the accuracy, maximize the comfort, and minimize the confusability of layout l (all metrics are normalized).

A.0.2 Speed. Speed was a quadratic term in our optimization process:

$$min \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{M} \sum_{l=1}^{M} p_{ij} t_{kl} x_{ik} x_{jl} \quad (2)$$

where x_{ik}, x_{jl} are binary decision variables indicating that a symbol $i(j)$ is associated with a nail region $k(l)$, p_{ij} is the frequency of the letter pair ij, and t_{kl} is the mean time it takes to touch nail region l after touching nail region k. We have a 26 letters (N) with 5 or 10 nail regions (M).

A.0.3 Accuracy. Accuracy was a second quadratic term:

$$max \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{M} \sum_{l=1}^{M} p_{ij} a_{kl} x_{ik} x_{jl} \quad (3)$$

Terms are identical to equation (2) save for a_{kl}, defined as the mean accuracy of touching nail region l after touching the nail region k.

A.0.4 Comfort. Comfort was a linear term:

$$max \sum_{i=1}^{N} \sum_{k=1}^{M} p_i c_k x_{ik} \quad (4)$$

where x_{ik} is the binary decision variable denoting whether or not a symbol i is mapped to an input action k, p_i the frequency of the symbol i, and c_i the comfort rating of touching the nail region k.

A.0.5 Confusability. Confusability was a third quadratic term:

$$min \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=1}^{M} (C_{ij} + C_{ji}) x_{ik} x_{jk} \quad (5)$$

where C is a confusability matrix [40] containing how frequently a symbol i is wrongly predicted to be a symbol j (for a given word prediction algorithm and data set) and x_{ij}, x_{jk} are binary decision variables that denote whether the two symbols are assigned to the same nail region (k)