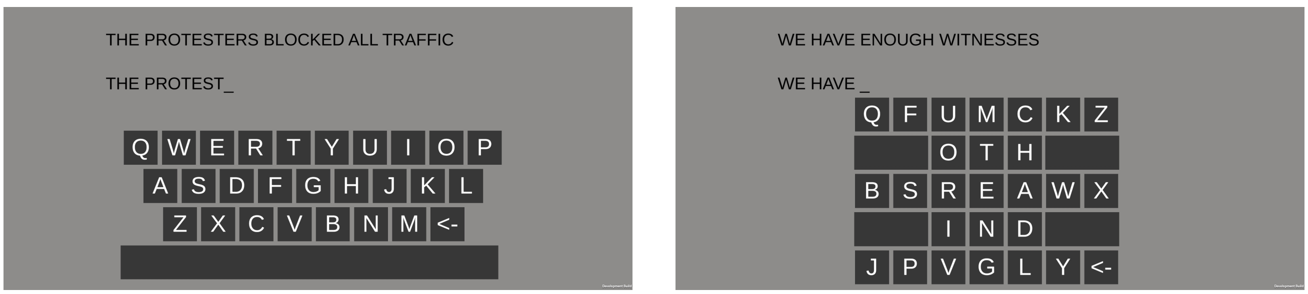
Does Repeatedly Typing the Same Phrase Provide a Good Estimate of Expert Text Entry Performance?

Aunnoy K Mutasim
amutasim@sfu.ca
School of Interactive Arts & Technology (SIAT)
Simon Fraser University
Vancouver, BC, Canada

Anil Ufuk Batmaz
ufuk.batmaz@concordia.ca
Computer Science & Software Engineering Department, Concordia University
Montreal, QC, Canada

Moaaz Hudhud Mughrabi
moaaz.mazen@stu.khas.edu.tr
Mechatronics Engineering Department, Kadir Has University
Istanbul, Turkey

Wolfgang Stuerzlinger
w.s@sfu.ca
School of Interactive Arts & Technology (SIAT)
Simon Fraser University
Vancouver, BC, Canada



(a) QWERTY

(b) OPTI

Figure 1: The two keyboard layouts.

ABSTRACT

To identify if novel/unfamiliar keyboard layouts like OPTI can outperform QWERTY, lengthy training through longitudinal studies is typically required. To reduce this logistical bottleneck, a popular approach in the literature requires participants to type the same phrase repeatedly. However, it is still unknown whether this approach provides a good estimate of expert performance. To validate this method, we set up a study where participants were tasked with typing the same phrase 96 times for both OPTI and QWERTY. Results showed that this approach has the potential to estimate expert performance for novel/unfamiliar keyboards faster than the traditional approach with different phrases. Yet, we also found that accurate estimates still require training over several days and, therefore, do not eliminate the need for a longitudinal study. Our findings thus show the need for research on faster, easier, and more reliable empirical approaches to evaluate text entry systems.

CCS CONCEPTS

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

KEYWORDS

Text Entry, Touch Typing, Tap Typing, Soft Keyboards, OPTI, QWERTY

ACM Reference Format:
Aunnoy K Mutasim, Anil Ufuk Batmaz, Moaaz Hudhud Mughrabi, and Wolfgang Stuerzlinger. 2023. Does Repeatedly Typing the Same Phrase Provide a Good Estimate of Expert Text Entry Performance?. In *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems (CHI EA '23), April 23–28, 2023, Hamburg, Germany*. ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/3544549.3585647

1 INTRODUCTION

Today, text entry has become an essential part of people's lives, e.g., for texting on smartphones, emailing, and writing reports and other documents. It also provides an alternative way to communicate for people with limited muscle control. [23, 35, 39, 42]. In most cases, text is entered using either a physical keyboard or a soft keyboard. Even though several more or less optimal layouts have been proposed [45], the QWERTY layout still dominates text entry to this day [13, 52].

CHI EA '23, April 23–28, 2023, Hamburg, Germany

ACM ISBN 978-1-4503-9422-2/23/04.
https://doi.org/10.1145/3544549.3585647

The primary reason for this is that the QWERTY layout has been widely available for over a century and therefore is habitually used in people's daily lives [52]. Thus, users are familiar with the QWERTY layout and the text entry speed they can achieve with it. Both facts make QWERTY also the most appropriate baseline for research on the evaluation of novel keyboard layouts. Previous work [31] has shown that theoretically more optimal keyboards can demonstrate significant benefits over QWERTY, but only after lengthy training. More specifically, users needed, on average, a total of 4 hours of practice with a novel keyboard layout, OPTI, to achieve the same text entry speed as with QWERTY. In the context of touch-based typing [31, 45], OPTI has been predicted to be about 30% faster than QWERTY; yet this point would only be reached after an average of 17 hours of training. Thus, it is not easy to convince regular smartphone and computer users to invest the effort associated with such extensive training, especially when the current typing speed of QWERTY is already deemed acceptable for most contexts.

Nevertheless, for interaction modalities/techniques that are more challenging to use, such as eye-gaze-based pointing and brain-computer interfaces for people with limited muscle control [23, 35, 39, 42], optimal layouts can potentially have a proportionally higher impact on text entry speeds and might thus be worth the effort of learning a new layout. Further, as discussed above, investigations of unfamiliar layouts suffer from a major methodological bottleneck, as demonstrating the expert performance potential of a given method requires that novices, i.e., first-time or beginner-level users of a system, be trained over a substantial amount of time, with some using up to 7.5 hours of each participant's time over 14 days, e.g., [10, 12, 13, 15, 19, 20, 22, 25–28, 31, 36, 38, 46, 53, 57]. Thus, running such studies is logistically expensive, especially when comparing novel/unfamiliar layouts to QWERTY [10, 31], or worse yet, when comparing several different layouts within the same study.

To reduce the logistical burden of a longitudinal study, some researchers have used an approach that requires participants to only type the same word(s)/phrase(s) repeatedly, e.g., [8, 9, 16, 54, 57]. Yet, this approach of estimating expert performance by repeatedly typing the same phrase, although popular in the literature [16, 54, 57], has not been validated before. More specifically, it is still unknown *if repeatedly typing the same phrase provides a good estimate of expert-level text entry performance* for a given keyboard layout. To address this gap in the literature, we investigated here whether this approach provides a good estimate of a layout's potential performance with expert users. For this, and similar to previous work [31], we compared OPTI and QWERTY by tasking participants to type the same phrase repeatedly. Our main contribution here is to show that repeatedly typing the same phrase approach does not accurately estimate expert-level text entry performance within a single day's training. Still, this approach potentially produces good estimates, but only through a (comparatively shorter) longitudinal study that still takes multiple days.

2 LITERATURE REVIEW

2.1 Evaluation Methods for Text Entry Systems

For any keyboard layout, a key difference between novices and experts is that experts can find keys faster as they are more familiar with the layout, incurring a (much) reduced average visual search time [32, 48]. Thus, and as mentioned above, text entry studies typically train novice users over a significant amount of time. One approach trained participants through multiple sessions on the same day, e.g., [13, 15, 22, 25–27, 38, 57]. Another, more externally valid approach is collecting data from trained users over several days, e.g., [10, 12, 19, 20, 28, 31, 36, 53], i.e., through a longitudinal study. See [21] for a comprehensive literature review of such studies.

One early longitudinal study was conducted by MacKenzie and Zhang [31], where the authors designed a theoretically more optimal novel keyboard interface, called OPTI, and compared it to QWERTY. Five participants typed by tapping the screen using a stylus on each of these keyboards. A single session was comprised of two typing rounds lasting 20-22 minutes each – one for each keyboard. The 20 sessions were separated by at least two hours but no more than two days. The average typing speed for OPTI increased from 17.0 words per minute (WPM) in session 1 to 44.3 WPM at the end. Similarly, QWERTY started at about 27.5 WPM and ended at about 40 WPM in session 20. Participants were able to type faster with OPTI starting from the 11$^{\text{th}}$ session, i.e., after about 4 hours of practice. Still, although participants experienced about 7 hours of practice with OPTI through the 20 sessions, this does not guarantee that they are experts. Thus, the authors then extrapolated the data using the power law of learning [20, 29] up to the 50$^{\text{th}}$ session and suggested Equation 1 and 2 for QWERTY and OPTI, respectively.

$$WPM_{QWERTY} = 27.597 \times Session^{0.1237}, R^2 = 0.9802 \quad (1)$$

$$WPM_{OPTI} = 17.24 \times Session^{0.3219}, R^2 = 0.9974 \quad (2)$$

In Equation 1 and 2, *Session* is the number of 20-22 minute training sessions, and R^2 is the squared correlation coefficient. These equations predict that after 17 hours of practice, QWERTY and OPTI have the potential to reach 44.8 WPM and 60.7 WPM, respectively. Thus, MacKenzie and Zhang [31] showed (in the context of touch-based keyboards) that theoretically more optimal layouts start to show significant benefits over QWERTY only after lengthy training.

In the domain of gaze-based keyboards, and to improve the speed of dwell-based systems, Majaranta et al. [36] conducted a longitudinal study where the participants were given the flexibility of adjusting the dwell time required to select a key. Eleven participants typed on a QWERTY keyboard for ten 15-minute sessions on ten separate days. Results showed that the typing speed increased from 6.9 WPM to 19.9 WPM in the final session. Also, the average dwell time decreased from 876 ms to 282 ms on average. However, due to the shorter dwell time, participants suffered more from the "Midas Touch" problem [55] as evident in the keystrokes per character (KSPC) results which increased from 1.09 in the first to 1.18 in the last session [36], as users needed more corrective actions to arrive at the correct result.

To bypass the burden of longitudinal studies, mathematical models of text entry performance (e.g., [20, 29, 32, 33, 49]) and empirical studies based on a system that simulates a perfect recognizer [22] have been used to predict expert performance. For touch-based text entry, Rick [45] developed a mathematical model for stroke-based (or swipe-based) typing based on Fitts' law using empirical data to predict expert-level performance on different keyboard layouts. Using this model, Rick [45] evaluated 22 existing layouts, e.g., QWERTY, Fitaly, Dvorak, OPTI I and II, Metropolis I and II, Quikwriting, and others, for both tapping and stroking/swiping. These layouts were also compared with two novel ones, named Square OSK and Hexagon OSK. These two layouts were generated based on the proposed model, i.e., optimized for swipe-based typing. The model predicted that stroking/swiping can achieve faster typing speeds compared to tapping for all 24 keyboard layouts. For example, stroking/swiping can achieve a 17.3% gain over tapping for QWERTY. Still, a much higher typing speed could be achieved if a "more suitable" layout is used, e.g., OPTI II is predicted to be 29.5% faster than QWERTY.

Magnien et al. [34] introduced an approach that provides visual clues to novices to help them speed up. After a user has typed a few characters, the system bolds the next few most likely characters, reducing the visual search space. Results showed that this approach can significantly reduce text entry time. Yet, this approach is not representative of expert behavior, as users still need to search for the next key in the middle of typing a word.

To avoid longitudinal studies, a popular approach employed by researchers requires participants to type the same word(s)/phrase(s) repeatedly, e.g., [8, 9, 16, 54, 57]. Although some of these studies [8, 9, 16, 57] also focused on investigating layout learning, they explicitly used the approach of repeatedly typing the same word(s)/phrase(s) to estimate expert performance. However, repeatedly typing a single phrase typically does not represent the character frequency of the used language, e.g., English, well, and thus, this approach cannot robustly evaluate real keyboard usage. To partially address this issue, Yu et al. [57] employed a slightly improved version where every participant was given a different phrase to repeatedly type 12 times. Still, Jokinen et al. [20] speculated that, although repeatedly typing the same word *"may"* perform better than QWERTY [8, 9], typing a phrase may not.

To validate Jokinen et al.'s [20] speculation about whether repeatedly typing the same phrase can actually predict expert-level performance, we replicated MacKenzie and Zhang's [31] study. Yet, instead of performing a longitudinal study and using different phrases, we tasked participants to type the same phrase 96 times over 8 sessions on the same day. We chose MacKenzie and Zhang's study [31] for two reasons. First, this choice allowed us to compare our results with the findings of a more externally valid longitudinal study and thus, helped validate the repeatedly typing the same phrase approach. Also, this study utilizes a touch-based input technique that still applies for smartphone-based text entry today, unlike, e.g., work on T9 [43].

2.2 Touch-Based Text Entry

Previous work introduced a variety of innovative text entry techniques for touch-based typing [1], including hand-posture adaptive keyboards [5, 17], tap-stroke hybrid keyboards [3], key-target resizing keyboards [18], gesture keyboards [2, 7, 37, 44, 59], and keyboards with different layouts [8, 14, 19, 58]. Yet, the focus of such studies is typically either on improving typing speeds or on novel text entry techniques. In contrast, we do not (*a priori*) aim to find a better touch-based text entry system. Instead, we take advantage of the ubiquity of touch-based systems to validate the approach of repeatedly typing the same phrase using MacKenzie and Zhang's [31] work.

3 USER STUDY

Here, we describe the details of our study and how we designed and evaluated the two keyboard layouts used by MacKenzie and Zhang [31], OPTI and QWERTY, with the approach of typing the same phrase repeatedly.

3.1 Apparatus and Keyboard Designs

Participants used their own smartphones to perform the experimental task since this allowed us to recruit remote participants easily and ensured that participants were using a device that they were most comfortable and experienced with from their day-to-day usage.

As our intent was to ensure comparability with MacKenzie and Zhang [31]'s work, we aimed to match their work as closely as possible. Thus, we purposely did not consider typing disambiguation methods, such as a model of the tap positions [45] or a language model [24], as this could confound the results.

Participants entered text with the OPTI and QWERTY layout in landscape mode. To be able to fit both layouts reasonably well onto the screen, we ensured that the devices had a display that was at least 6" diagonally. To fit the keyboard into such a 6" screen, we still had to set the key sizes to be 0.2 cm smaller than the original design by MacKenzie and Zhang [31], i.e., the keys were 0.8×0.8 cm. As suggested by Fitts' law [6, 11, 40, 41, 47] and previous work [32, 56], this slight decrease in size should not significantly impact the results. We also introduced a gap of 0.1 cm between keys to avoid unintended selections when the user touches the edge between two keys and provided auditory feedback through a subtle click for each key 'hit'. Instead of the 'F1' key of the original OPTI layout, and just like any smartphone QWERTY keyboard layout, we added a backspace key beside the character 'M' for QWERTY to support error correction. The Android app for the study was developed using Unity. Figure 1 shows the design of the two keyboard layouts.

3.2 Participants and Procedure

Eight participants (7 male), aged 31.4 ± 5.21 years, took part in the study. They were recruited via word of mouth and ads over social media platforms and email and were compensated with $15 for their participation. All participants had over nine years of experience with typing on a physical/soft QWERTY keyboard, but none had experience with the OPTI keyboard prior to this study.

Participants started by filling out a consent form and a demographic questionnaire which included questions regarding their age, gender, and experience using OPTI and QWERTY layouts. Then, they typed using both OPTI and QWERTY, presented in counterbalanced order, i.e., we used a within-subjects experimental design with just one independent variable – the two keyboard layouts. Each participant typed the same phrase 96 times in eight sessions (i.e., 12 times per session) for each of the two layouts. Participants were instructed to type in an extra space once they were done typing a phrase to clearly denote the end of that phrase. Participants rested for 2 minutes between sessions. The chosen phrase was selected randomly from MacKenzie and Soukoreff's set [30] (comprises 500 phrases in total with the minimum, maximum, and average phrase lengths being 16, 43, and 28.6 characters, respectively). We followed Yu et al.'s approach [57] and ensured that every participant was given a different/unique phrase to transcribe to increase the validity.

Instead of a stylus [31], participants used the index finger of their dominant hand to type. They were also asked to correct mistakes immediately but to ignore mistakes that occurred two or more letters back, simply to avoid too many corrections, i.e., to avoid affecting their typing speed too much. The task was performed sitting in a chair, and participants held the phone with their non-dominant hand. Between the two keyboard conditions, participants rested for at least 5 minutes. At the end of the typing task, we conducted a short semi-structured interview where participants shared their typing experience with the two keyboard layouts. In total, the study took about an hour per participant.

3.3 Performance Metrics

We used the following metrics to evaluate text entry performance:

- *Words per minute (WPM)*, which is the total number of words typed per minute, where a single word comprises a sequence of any 5 characters including spaces [4].
- *Keystrokes per character (KSPC)*, which is the ratio of the number of keys selected to the length of the typed text [50]. In other words, KSPC is the number of key selections required to (correctly) type a single character, including the extra keystrokes required for error correction, i.e., when the backspace key was hit.
- *Minimum String Distance Error Rate (MSD ER)*, where MSD is the minimum number of insertions, deletions, and substitutions required to transform one string into another. We use the formulation of the MSD ER metric introduced by Soukoreff and MacKenzie [51]. Unlike MacKenzie and Zhang's study [31] where only the error rate was reported, which suffers from limitations (see [50, 51]), we chose to report the MSD ER metric.

4 RESULTS AND DISCUSSION

The results of the study are shown in Table 1 and Figure 2. As can be seen in Figure 2a, although repeatedly typing the same phrase improved participants' performance for OPTI, the data for OPTI never approached the performance with QWERTY. In other words, participants' typing speed using the OPTI layout was always below the typing speed achieved using QWERTY throughout the eight sessions.

As per Table 1 and Figure 2a, participants started at an average typing speed of 19.1 ± 5.85 WPM in the first and ended with 31.0 ± 8.64 WPM in the last session with OPTI, which was also the fastest average typing speed across participants. The fastest individual participant typed at an average speed of 41.1 ± 5.12 WPM on the 8^{th} session with OPTI. Participants achieved 33.7 ± 6.37 WPM on the first and 33.0 ± 5.94 WPM on the last session using QWERTY. The highest typing speed of 36.7 ± 6.15 WPM for QWERTY was achieved by the participants on the 4^{th} session (see Table 1). The fastest participant was able to reach an average typing speed of 42.6 ± 7.87 WPM on the 4^{th} session with QWERTY.

For both keyboard layouts, we did not observe noteworthy trends in terms of KSPC and MSD ER. Other than a few exceptions, i.e., sessions 1 and 3-5 in Figure 2b, and sessions 3 and 8 in Figure 2c, the curves for both keyboards seemed to be quite similar to each other. Also, average MSD ER for both OPTI and QWERTY were all below 2% throughout the eight sessions (see Figure 2c), showing that the participants were quite careful on both the keyboards layouts [24].

When referencing our results with Equation 1 and 2 [31], the average typing speed of the last session is equivalent to 4.3 (20-22 minute) sessions for QWERTY and 6.2 sessions for OPTI (see Table 1). For QWERTY, where the fastest session was not the last, we estimate that it would take 10 sessions to reach speeds of 36.7 WPM, as achieved in the 4^{th} session. This shows that repeatedly typing the same phrase for about 30 minutes can – at best – achieve comparable results to 3.67 hours of normal/traditional training (i.e., training with different phrases) with a known, i.e., QWERTY, or 2.27 hours of traditional training with an unknown keyboard layout, in this case, OPTI.

In contrast to our work, MacKenzie and Zhang [31] found that OPTI outperformed QWERTY starting from the 11^{th} (20-22 minute) session and eventually reached 44.3 WPM in session 20, with QWERTY achieving 40 WPM on the same session. The typing speed of our participants did not even get close to the expert performance reported by MacKenzie and Zhang [31] for either keyboard. The results of our study thus indicate that – although repeatedly typing the same phrase improves performance – this approach is not suitable for reliably estimating expert performance for unfamiliar keyboard layouts. Thus, our results validate Jokinen et al.'s [20] speculation that repeatedly typing a phrase on an unfamiliar layout like OPTI does not demonstrate that it can surpass QWERTY's performance, at least not with a single day's training. More importantly, *the approach of repeatedly typing the same phrase does not come close to a good estimate of expert-level text entry performance with a single day's training.*

The change in typing speed over time for OPTI in Figure 2a showed an increasing trend until session 6. From there onwards, the curve completely flattened out. This is evidence that running more sessions probably would not have increased OPTI's performance any further. As for the trend of the curve for QWERTY in Figure 2a, the typing speed slightly increased from the 1^{st} to the 4^{th} session. However, a decreasing trend can be observed from the 5^{th} session onwards.

One participant in the semi-structured interview mentioned that *"the task very quickly got boring and frustrating as it seemed like it was never going to end."* Another explained *"My mind kept wandering off. It was very hard to continuously keep my concentration on the task."* Similarly, another participant shared *"After a couple of sessions, QWERTY was especially hard as there was no challenge associated with the task, ... , unlike OPTI, where I felt there was still scope for improvement."* Others also shared similar feedback about the experimental task. Given these insights, we believe that the downward trend of QWERTY's typing speed starting from the 5^{th} session is directly associated with the fatigue of the participants and the lack of challenge in the experimental task. This opens up

Table 1: WPM results for each session and using Equation 1 and 2, the corresponding number of (22 minute) Training Sessions and Training Time required to reach that WPM when different phrases are typed for training.

	Typing Speed (WPM)		Projected No. of Training Sessions		Projected Training Time (hours)	
Session	OPTI	QWERTY	OPTI	QWERTY	OPTI	QWERTY
1	19.1 ± 5.85	33.7 ± 6.37	1.4	5.0	0.50	1.83
2	22.9 ± 7.01	36.0 ± 6.35	2.4	8.6	0.89	3.14
3	25.4 ± 7.11	35.1 ± 6.12	3.3	7.0	1.23	2.56
4	25.7 ± 7.24	36.7 ± 6.15	3.5	10.0	1.27	3.67
5	27.9 ± 8.08	35.5 ± 5.76	4.5	7.6	1.64	2.80
6	31.0 ± 7.97	34.4 ± 7.73	6.2	5.9	2.28	2.16
7	30.8 ± 7.86	34.1 ± 5.6	6.1	5.5	2.22	2.03
8	31.0 ± 8.64	33.0 ± 5.94	6.2	4.3	2.27	1.57

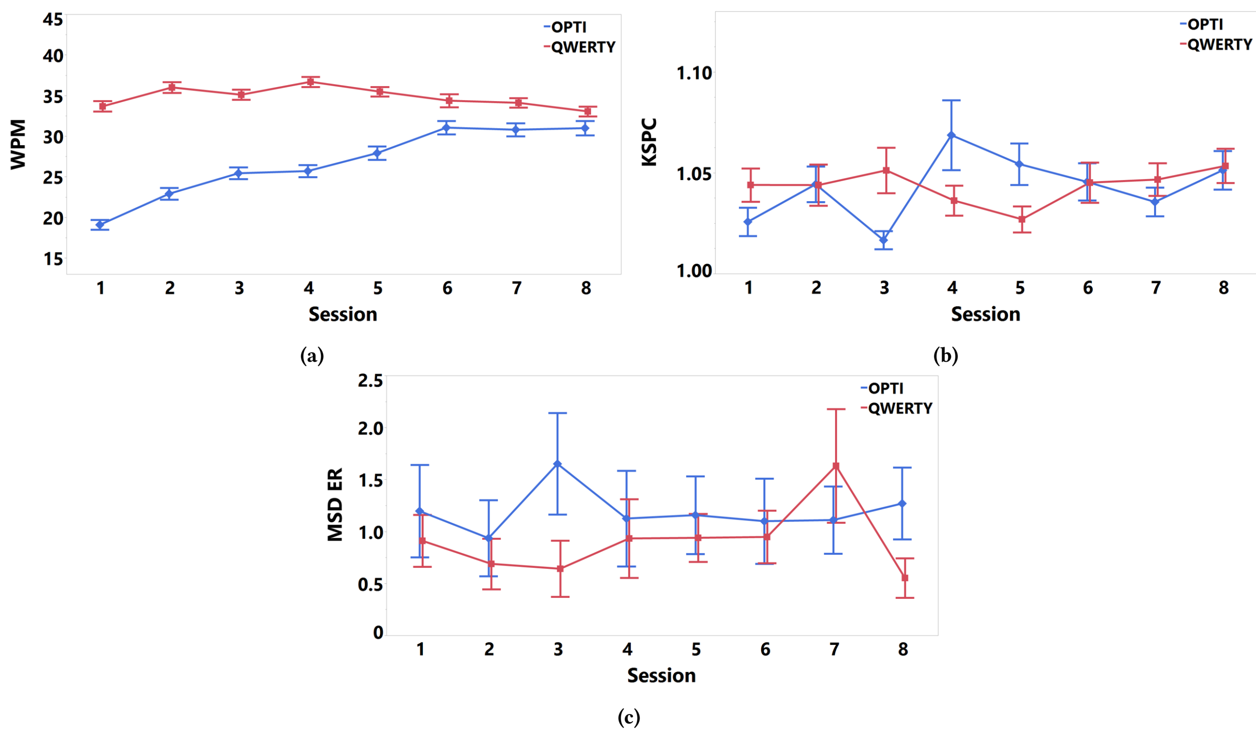


Figure 2: (a) WPM, (b) KSPC, and (c) MSD ER results over eight sessions of typing the same phrase. The error bars show the standard error of means.

the question of whether this fatigue impacted the results for the flattening trend of OPTI starting from the 6th session as well.

To analyze this we fit a regression based on the power law of learning through the WPM data [20, 29, 31], see Figure 3, which gave us the following two equations:

$$WPM_{QWERTY} = 35.684 \times Session^{-0.016}, R^2 = 0.0517 \quad (3)$$

$$WPM_{OPTI} = 15.64 \times Session^{0.3273}, R^2 = 0.9596 \quad (4)$$

In Equation 3 and 4, a single *Session* is comprised of repeatedly typing the same phrase 12 times, and R^2 represents the squared correlation coefficient. As per the R^2 value in Equation 4, approximately 96% of the variance is accounted for in the fitted learning model, which means that the model is a very good prediction of user behavior. Equation 4 predicts that it would require ≈ 24 sessions (289 repetitions, 8 sessions per day = 4 days) of typing the same phrase repeatedly with OPTI before participants can reach the typing speed of 44.3 WPM which was reported by MacKenzie and Zhang [31] to take twenty (20-22 minute) normal/traditional (i.e., training with different phrases) training sessions (i.e., 20 days) of training. In other words, although repeatedly typing the same phrase is predicted to be able to provide a good estimate of expert performance and can do so faster than traditional training, this approach would still require training participants over several days and would therefore **not** eliminate the need for a longitudinal study.

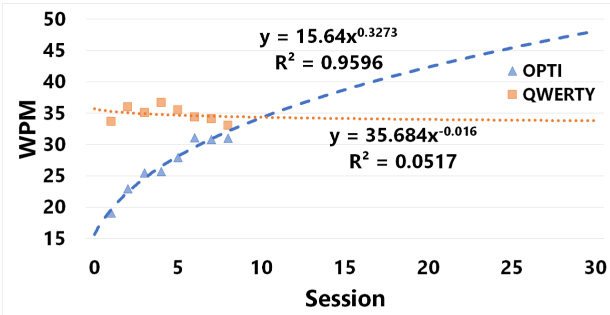


Figure 3: WPM over eight sessions along with an extrapolation of the learning curve to the 30th session.

Yet, repeatedly typing the same phrase on a known layout, i.e., QWERTY, had detrimental effects on learning, as evident by the negative exponent (i.e., -0.016) and a very small R^2 value in Equation 3. Therefore, we suggest avoiding studies that involve repeatedly typing the same phrase task for more than 4 sessions, i.e., 48 repetitions, for QWERTY or any layout already very well known by a participant. Overall, our analysis based on the power law of learning showed that the observed flattening trend in typing speed of OPTI from the 6th session and the decreasing trend of QWERTY from 5th session onwards are most likely due to the fatigue associated with the experimental task.

Researchers had previously used the repetitive phrase typing approach either with complete novices, e.g., [54], or with users who were first trained for a small number of sessions using the traditional approach of typing different phrases and then typing the same phrase for a few more sessions, e.g., [9, 16, 57]. In our study, we chose the first method [54], which uses only a single phrase repetitively, as we wanted to investigate an approach that might reach expert performance faster. Further, we ensured that the total number of repetitions performed by participants in our study was a lot larger than the combination of phrases and repetitions typically employed by other studies [9, 57].

5 CONCLUSION AND FUTURE WORK

Here we experimented with the popular approach of repeatedly typing the same phrase and investigated if it provides a good estimate of a keyboard layout's potential performance with expert users. Results showed that this approach has the potential to estimate expert performance for novel/unfamiliar keyboards; however, only after lengthy training spanning several days. Thus, although shorter than compared to the traditional approach of training users with different phrases, a longitudinal study is still inevitable. We also found that repeating the typing task too many times for a known keyboard layout, like QWERTY, has detrimental effects on the results as participants tend to lose concentration and get mentally tired faster than with an unfamiliar keyboard layout. We believe that *our findings demonstrate the need for further investigation on finding a faster, easier, and more reliable empirical approach to evaluate text entry systems.*

In the future, we plan to investigate how repeatedly typing the same word compares to repeatedly typing the same phrase and the traditional approach of training users with different phrases/words. Also, it would be interesting to investigate in the future how the two approaches to repetitive phrase typing, i.e., with or without prior training using the traditional approach, relate to each other. Finally, although using different phrases for different participants for the repetitive typing task increases external validity, the chosen phrases could potentially have a confounding effect on the results, which also requires further research.

REFERENCES

[1] Ohoud Alharbi, Ahmed Sabbir Arif, Wolfgang Stuerzlinger, Mark D. Dunlop, and Andreas Komninos. 2019. WiseType : a tablet keyboard with color-coded visualization and various editing options for error correction. In *Graphics Interface 2019*. CHCCS, CAN, 10 pages. https://strathprints.strath.ac.uk/67396/

[2] Ouais Alsharif, Tom Ouyang, Françoise Beaufays, Shumin Zhai, Thomas Breuel, and Johan Schalkwyk. 2015. Long short term memory neural network for keyboard gesture decoding. In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2076–2080. https://doi.org/10.1109/ICASSP.2015.7178336

[3] A.S. Arif, Michel Pahud, K. Hinckley, and B. Buxton. 2014. Experimental study of stroke shortcuts for a touchscreen keyboard with gesture-redundant keys removed. In *Proceedings - Graphics Interface*. CHCCS, 43–50.

[4] A. S. Arif and W. Stuerzlinger. 2009. Analysis of text entry performance metrics. In *2009 IEEE Toronto International Conference Science and Technology for Humanity (TIC-STH)*. IEEE, 100–105. https://doi.org/10.1109/TIC-STH.2009.5444533

[5] Shiri Azenkot and Shumin Zhai. 2012. Touch Behavior with Different Postures on Soft Smartphone Keyboards. In *Proceedings of the 14th International Conference on Human-Computer Interaction with Mobile Devices and Services* (San Francisco, California, USA) *(MobileHCI '12)*. Association for Computing Machinery, New York, NY, USA, 251–260. https://doi.org/10.1145/2371574.2371612

[6] Anil Ufuk Batmaz, Aunnoy K Mutasim, Morteza Malekmakan, Elham Sadr, and Wolfgang Stuerzlinger. 2020. Touch the Wall: Comparison of Virtual and Augmented Reality with Conventional 2D Screen Eye-Hand Coordination Training Systems. In *2020 IEEE Conference on Virtual Reality and 3D User Interfaces (VR)*. IEEE, 184–193. https://doi.org/10.1109/VR46266.2020.00037

[7] 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

[8] 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

[9] 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

[10] Steven J. Castellucci, I. Scott MacKenzie, Mudit Misra, Laxmi Pandey, and Ahmed Sabbir Arif. 2019. TiltWriter: Design and Evaluation of a No-Touch Tilt-Based Text Entry Method for Handheld Devices. In *Proceedings of the 18th International Conference on Mobile and Ubiquitous Multimedia* (Pisa, Italy) *(MUM '19)*. Association for Computing Machinery, New York, NY, USA, Article 7, 8 pages. https://doi.org/10.1145/3365610.3365629

[11] Mungyeong Choe, Yeongcheol Choi, Jaehyun Park, and Hyun K. Kim. 2019. Comparison of Gaze Cursor Input Methods for Virtual Reality Devices. *International Journal of Human–Computer Interaction* 35, 7 (2019), 620–629. https://doi.org/10.1080/10447318.2018.1484054 arXiv:https://doi.org/10.1080/10447318.2018.1484054

[12] Edward Clarkson, James Clawson, Kent Lyons, and Thad Starner. 2005. An Empirical Study of Typing Rates on Mini-QWERTY Keyboards. In *CHI '05 Extended Abstracts on Human Factors in Computing Systems* (Portland, OR, USA) *(CHI EA '05)*. Association for Computing Machinery, New York, NY, USA, 1288–1291. https://doi.org/10.1145/1056808.1056898

[13] Tafadzwa Joseph Dube and Ahmed Sabbir Arif. 2019. Text entry in virtual reality: A comprehensive review of the literature. In *International Conference on Human-Computer Interaction*. Springer, 419–437.

[14] Mark Dunlop and John Levine. 2012. Multidimensional Pareto Optimization of Touchscreen Keyboards for Speed, Familiarity and Improved Spell Checking. 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, 2669–2678. https://doi.org/10.1145/2207676.2208659

[15] Wenxin Feng, Jiangnan Zou, Andrew Kurauchi, Carlos H Morimoto, and Margrit Betke. 2021. HGaze Typing: Head-Gesture Assisted Gaze Typing. In *ACM Symposium on Eye Tracking Research and Applications*. Association for Computing Machinery, New York, NY, USA, Article 11, 11 pages. https://doi.org/10.1145/3448017.3457379

[16] Sanjay Ghosh, Anirudha Joshi, Manjiri Joshi, Nagraj Emmadi, Girish Dalvi, Shashank Ahire, and Swati Rangale. 2017. Shift+Tap or Tap+LongPress? The Upper Bound of Typing Speed on InScript. 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, 2059–2063. https://doi.org/10.1145/3025453.3025944

[17] Mayank Goel, Alex Jansen, Travis Mandel, Shwetak N. Patel, and Jacob O. Wobbrock. 2013. ContextType: Using Hand Posture Information to Improve Mobile Touch Screen Text Entry. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Paris, France) *(CHI '13)*. Association for Computing Machinery, New York, NY, USA, 2795–2798. https://doi.org/10.1145/2470654.2481386

[18] Asela Gunawardana, Tim Paek, and Christopher Meek. 2010. Usability Guided Key-Target Resizing for Soft Keyboards. In *Proceedings of the 15th International Conference on Intelligent User Interfaces* (Hong Kong, China) *(IUI '10)*. Association for Computing Machinery, New York, NY, USA, 111–118. https://doi.org/10.1145/1719970.1719986

[19] Hiroyuki Hakoda, Buntarou Shizuki, and Jiro Tanaka. 2016. QAZ keyboard: QWERTY based portrait soft keyboard. In *International Conference of Design, User Experience, and Usability*. Springer, 24–35.

[20] Jussi P. P. Jokinen, Sayan Sarcar, Antti Oulasvirta, Chaklam Silpasuwanchai, Zhenxin Wang, and Xiangshi Ren. 2017. Modelling Learning of New Keyboard Layouts. 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, 4203–4215. https://doi.org/10.1145/3025453.3025580

[21] Maria Kjærup, Mikael B. Skov, Peter Axel Nielsen, Jesper Kjeldskov, Jens Gerken, and Harald Reiterer. 2021. *Longitudinal Studies in HCI Research: A Review of CHI Publications From 1982–2019*. Springer International Publishing, Cham, 11–39. https://doi.org/10.1007/978-3-030-67322-2_2

[22] Per Ola Kristensson and Keith Vertanen. 2012. The Potential of Dwell-Free Eye-Typing for Fast Assistive Gaze Communication. In *Proceedings of the Symposium on Eye Tracking Research and Applications* (Santa Barbara, California) *(ETRA '12)*. Association for Computing Machinery, New York, NY, USA, 241–244. https://doi.org/10.1145/2168556.2168605

[23] Chandan Kumar, Raphael Menges, Daniel Müller, and Steffen Staab. 2017. Chromium Based Framework to Include Gaze Interaction in Web Browser. In *Proceedings of the 26th International Conference on World Wide Web Companion* (Perth, Australia) *(WWW '17 Companion)*. International World Wide Web Conferences Steering Committee, Republic and Canton of Geneva, CHE, 219–223. https://doi.org/10.1145/3041021.3054730

[24] Andrew Kurauchi, Wenxin Feng, Ajjen Joshi, Carlos Morimoto, and Margrit Betke. 2016. EyeSwipe: Dwell-Free Text Entry Using Gaze Paths. 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, 1952–1956. https://doi.org/10.1145/2858036.2858335

[25] Andrew Kurauchi, Wenxin Feng, Ajjen Joshi, Carlos H. Morimoto, and Margrit Betke. 2020. Swipe&Switch: Text Entry Using Gaze Paths and Context Switching. In *Adjunct Publication of the 33rd Annual ACM Symposium on User Interface Software and Technology* (Virtual Event, USA) *(UIST '20 Adjunct)*. Association for Computing Machinery, New York, NY, USA, 84–86. https://doi.org/10.1145/3379350.3416193

[26] Frank Chun Yat Li, Richard T. Guy, Koji Yatani, and Khai N. Truong. 2011. The 1line Keyboard: A QWERTY Layout in a Single Line. In *Proceedings of the 24th Annual ACM Symposium on User Interface Software and Technology* (Santa Barbara, California, USA) *(UIST '11)*. Association for Computing Machinery, New York, NY, USA, 461–470. https://doi.org/10.1145/2047196.2047257

[27] X. Lu, D. Yu, H. N. Liang, W. Xu, Y. Chen, X. Li, and K. Hasan. 2020. Exploration of Hands-free Text Entry Techniques For Virtual Reality. In *2020 IEEE International Symposium on Mixed and Augmented Reality (ISMAR)*. IEEE, 344–349. https://doi.org/10.1109/ISMAR50242.2020.00061

[28] Kent Lyons, Thad Starner, Daniel Plaisted, James Fusia, Amanda Lyons, Aaron Drew, and E. W. Looney. 2004. Twiddler Typing: One-Handed Chording Text Entry for Mobile Phones. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vienna, Austria) *(CHI '04)*. Association for Computing Machinery, New York, NY, USA, 671–678. https://doi.org/10.1145/985692.985777

[29] I. Scott MacKenzie and R. William Soukoreff. 2002. Text Entry for Mobile Computing: Models and Methods,Theory and Practice. *Human–Computer Interaction* 17, 2-3 (2002), 147–198. https://doi.org/10.1080/07370024.2002.9667313 arXiv:https://www.tandfonline.com/doi/pdf/10.1080/07370024.2002.9667313

[30] I. Scott MacKenzie and R. William Soukoreff. 2003. Phrase Sets for Evaluating Text Entry Techniques. In *CHI '03 Extended Abstracts on Human Factors in Computing Systems* (Ft. Lauderdale, Florida, USA) *(CHI EA '03)*. Association for Computing Machinery, New York, NY, USA, 754–755. https://doi.org/10.1145/765891.765971

[31] 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

[32] I. Scott MacKenzie and Shawn X. Zhang. 2001. An empirical investigation of the novice experience with soft keyboards. *Behaviour & Information Technology* 20, 6 (2001), 411–418. https://doi.org/10.1080/01449290110089561 arXiv:https://doi.org/10.1080/01449290110089561

[33] I. Scott Mackenzie, Shawn X. Zhang, and R. William Soukoreff. 1999. Text entry using soft keyboards. *Behaviour & Information Technology* 18, 4 (1999), 235–244. https://doi.org/10.1080/014492999118995 arXiv:https://doi.org/10.1080/014492999118995

[34] Laurent Magnien, Jean Léon Bouraoui, and Nadine Vigouroux. 2004. Mobile Text Input with Soft Keyboards: Optimization by Means of Visual Clues. In *Mobile Human-Computer Interaction - MobileHCI 2004*, Stephen Brewster and Mark Dunlop (Eds.). Springer Berlin Heidelberg, Berlin, Heidelberg, 337–341.

[35] Päivi Majaranta. 2012. Communication and text entry by gaze. In *Gaze interaction and applications of eye tracking: Advances in assistive technologies*. IGI Global, 63–77.

[36] Päivi Majaranta, Ulla-Kaija Ahola, and Oleg Špakov. 2009. Fast Gaze Typing with an Adjustable Dwell Time. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Boston, MA, USA) *(CHI '09)*. Association for Computing Machinery, New York, NY, USA, 357–360. https://doi.org/10.1145/1518701.1518758

[37] Anders Markussen, Mikkel Rønne Jakobsen, and Kasper Hornbæk. 2014. Vulture: A Mid-Air Word-Gesture Keyboard. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Toronto, Ontario, Canada) *(CHI '14)*. Association for Computing Machinery, New York, NY, USA, 1073–1082. https://doi.org/10.1145/2556288.2556964

[38] Carlos H. Morimoto, Jose A. T. Leyva, and Antonio Diaz-Tula. 2018. Context Switching Eye Typing Using Dynamic Expanding Targets. In *Proceedings of the Workshop on Communication by Gaze Interaction* (Warsaw, Poland) *(COGAIN '18)*. Association for Computing Machinery, New York, NY, USA, Article 6, 9 pages. https://doi.org/10.1145/3206343.3206347

[39] Martez E. Mott, Shane Williams, Jacob O. Wobbrock, and Meredith Ringel Morris. 2017. Improving Dwell-Based Gaze Typing with Dynamic, Cascading Dwell Times. 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, 2558–2570. https://doi.org/10.1145/3025453.3025517

[40] Aunnoy Mutasim, Anil Ufuk Batmaz, Moaaz Hudhud Mughrabi, and Wolfgang Stuerzlinger. 2022. Performance Analysis of Saccades for Primary and Confirmatory Target Selection. In *Proceedings of the 28th ACM Symposium on Virtual Reality Software and Technology* (Tsukuba, Japan) *(VRST '22)*. Association for Computing Machinery, New York, NY, USA, Article 18, 12 pages. https://doi.org/10.1145/3562939.3565619

[41] Aunnoy K Mutasim, Anil Ufuk Batmaz, and Wolfgang Stuerzlinger. 2021. Pinch, Click, or Dwell: Comparing Different Selection Techniques for Eye-Gaze-Based Pointing in Virtual Reality. In *ACM Symposium on Eye Tracking Research and Applications* (Virtual Event, Germany) *(ETRA '21 Short Papers)*. Association for Computing Machinery, New York, NY, USA, Article 15, 7 pages. https://doi.org/10.1145/3448018.3457998

[42] Chethan Pandarinath, Paul Nuyujukian, Christine H Blabe, Brittany L Sorice, Jad Saab, Francis R Willett, Leigh R Hochberg, Krishna V Shenoy, and Jaimie M Henderson. 2017. High performance communication by people with paralysis using an intracortical brain-computer interface. *Elife* 6 (2017), e18554.

[43] Andriy Pavlovych and Wolfgang Stuerzlinger. 2004. Model for Non-Expert Text Entry Speed on 12-Button Phone Keypads. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Vienna, Austria) *(CHI '04)*. Association for Computing Machinery, New York, NY, USA, 351–358. https://doi.org/10.1145/985692.985737

[44] Shyam Reyal, Shumin Zhai, and Per Ola Kristensson. 2015. Performance and User Experience of Touchscreen and Gesture Keyboards in a Lab Setting and in the Wild. 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, 679–688. https://doi.org/10.1145/2702123.2702597

[45] Jochen Rick. 2010. Performance Optimizations of Virtual Keyboards for Stroke-Based Text Entry on a Touch-Based Tabletop. In *Proceedings of the 23nd Annual ACM Symposium on User Interface Software and Technology* (New York, New York, USA) *(UIST '10)*. Association for Computing Machinery, New York, NY, USA, 77–86. https://doi.org/10.1145/1866029.1866043

[46] Daniel Rough, Keith Vertanen, and Per Ola Kristensson. 2014. An Evaluation of Dasher with a High-Performance Language Model as a Gaze Communication Method. In *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces* (Como, Italy) *(AVI '14)*. Association for Computing Machinery, New York, NY, USA, 169–176. https://doi.org/10.1145/2598153.2598157

[47] Immo Schuetz, T. Scott Murdison, Kevin J. MacKenzie, and Marina Zannoli. 2019. An Explanation of Fitts' Law-like Performance in Gaze-Based Selection Tasks Using a Psychophysics Approach. 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, 1–13. https://doi.org/10.1145/3290605.3300765

[48] Andrew Sears, Julie A. Jacko, Josey Chu, and Francisco Moro. 2001. The role of visual search in the design of effective soft keyboards. *Behaviour & Information Technology* 20, 3 (2001), 159–166. https://doi.org/10.1080/01449290110049790 arXiv:https://doi.org/10.1080/01449290110049790

[49] 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

[50] R. William Soukoreff and I. Scott MacKenzie. 2001. Measuring Errors in Text Entry Tasks: An Application of the Levenshtein String Distance Statistic. In *CHI '01 Extended Abstracts on Human Factors in Computing Systems* (Seattle, Washington) *(CHI EA '01)*. Association for Computing Machinery, New York, NY, USA, 319–320. https://doi.org/10.1145/634067.634256

[51] R. William Soukoreff and I. Scott MacKenzie. 2003. Metrics for Text Entry Research: An Evaluation of MSD and KSPC, and a New Unified Error Metric. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (Ft. Lauderdale, Florida, USA) *(CHI '03)*. Association for Computing Machinery, New York, NY, USA, 113–120. https://doi.org/10.1145/642611.642632

[52] Marco Speicher, Anna Maria Feit, Pascal Ziegler, and Antonio Krüger. 2018. Selection-Based Text Entry in Virtual Reality. 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/10.1145/3173574.3174221

[53] Outi Tuisku, Päivi Majaranta, Poika Isokoski, and Kari-Jouko Räihä. 2008. Now Dasher! Dash Away! Longitudinal Study of Fast Text Entry by Eye Gaze. In *Proceedings of the 2008 Symposium on Eye Tracking Research & Applications* (Savannah, Georgia) *(ETRA '08)*. Association for Computing Machinery, New York, NY, USA, 19–26. https://doi.org/10.1145/1344471.1344476

[54] Mario H. Urbina and Anke Huckauf. 2007. Dwell time free eye typing approaches. In *Proceedings of the 3rd Conference on Communication by Gaze Interaction* (Leicester, UK) *(COGAIN '07)*. Association for Computing Machinery, New York, NY, USA, 65–70.

[55] Boris Velichkovsky, Andreas Sprenger, and Pieter Unema. 1997. Towards gaze-mediated interaction: Collecting solutions of the "Midas touch problem". In *Human-Computer Interaction INTERACT'97*. Springer, 509–516.

[56] Keith Vertanen, Haythem Memmi, Justin Emge, Shyam Reyal, and Per Ola Kristensson. 2015. VelociTap: Investigating Fast Mobile Text Entry Using Sentence-Based Decoding of Touchscreen Keyboard Input. 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, 659–668. https://doi.org/10.1145/2702123.2702135

[57] 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

[58] 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

[59] Shumin Zhai and Per Ola Kristensson. 2012. The Word-Gesture Keyboard: Reimagining Keyboard Interaction. *Commun. ACM* 55, 9 (sep 2012), 91–101. https://doi.org/10.1145/2330667.2330689