Is Touch-based Text Input Practical for a Smartwatch?

Barbara S. Chaparro[(✉)], Jibo He, Colton Turner, and Kirsten Turner

Wichita State University, Wichita, KS, USA
{barbara.chaparro,jibo.he,cjturner,
kxturner}@wichita.edu

Abstract. The use of smartwatches is increasing exponentially as is consumer interest. Currently, smartwatches offer the ability to read text messages, notifications, and email once they are synchronized with a smartphone. Text input, however, is limited to voice or predefined response phrases and no input keyboard is typically provided. A general consensus is that the interface of a smartwatch may be too small to implement a QWERTY keyboard. This study examined user performance and acceptance with two commercially available QWERTY keyboards, Swype and Fleksy, on a smartwatch. Contrary to the suspicion about the small screen of a smartwatch for text input, results indicate users can type accurately at speeds averaging 20–30 words per minute after brief practice, which is comparable to the typing speed of novice smartphone users.

Keywords: Smartwatch · Text input · Fleksy · Swype · Touch · Wearable

1 Introduction

Recent reports estimate over 15 million smartwatches will be sold in 2015 [1]. Experts project this number to increase to 90 million smartwatches globally in 2018 [2] and 300 million by 2020 [1]. Smartwatches let users check text messages, notifications, and emails without having to look at their phone. Responses to text notifications on smartwatches, however, are limited to voice input or a choice of predefined text responses.

The use of a small QWERTY keyboard for text entry on a smartwatch has not been implemented until recently (Samsung Gear S). Presumably this is because of the limited space for a full-size keyboard and limited touch sensitive zones given the average size of the human finger [3, 4]. Instead, many research teams have developed alternative keyboards and text input methods for such small displays. These solutions include a touch-sensitive wristband [5], touch-sensitive zooming of keys [6], round touchscreen key entry [7], panning and twist action input [8], quasi-QWERTY design [9] and a one-key QWERTY [10]. Typing performance of these alternative input methods ranges from 5–10 words per minute (WPM), along with a learning curve for new users. Text input performance with the only traditional QWERTY solution (one-key keyboard) was found to be as high as 24 WPM after 5 20-minute sessions. However, the keyboard size was 70 × 35 mm, which is approximately twice that of today's smartwatch display [10].

Funk, Sahami, Henze, and Schmidt [5] developed criteria for smartwatch text input based on the literature and their observations. The criteria include that (1) users will interact using finger input (instead of stylus) (2) the watch face will be a touchscreen (3) there will be limited dynamic feedback (4) users are likely to input short length text (5) the watch must have editing support (6) target key size must be larger than 7 mm, and (7) the watch must support some gesture-based control. Preliminary results with a new alphanumeric keyboard showed that users could learn it fairly quickly, but performance was compromised by the screen sensitivity and layout problems.

Several smartphone text input methods are available today that rely on sophisticated algorithms to predict a user's intended input. Two such systems are Swype (a trace-based input method) and Fleksy (a tap and gesture based method). Both systems use a familiar QWERTY-style keyboard and are available for Android smartphone devices, but are not yet provided on smartwatches out of the box.

2 Purpose

This study tests the feasibility of typing on a smartwatch by comparing typing performance and user satisfaction of two text input methods that use a standard QWERTY keyboard.

3 Method

3.1 Participants

Eighteen volunteers (12 Females and 6 Males) recruited from a midwestern University, ranging in age from 18 to 42 years old (M = 22.27, SD = 7.67), participated in this study. All participants were fluent English speakers, had normal or corrected-to-normal vision, owned a touch screen smartphone, and did not have any physical limitations to their hands that would prevent them from being able to text. All participants had prior typing experience on a touch screen phone. No participants had prior experience typing on a smartwatch. Participants received course credit for their participation.

3.2 Materials

A Samsung Galaxy Gear 1 smartwatch with a display size of 1.62" (41.40 mm) was used in this study. Two input methods, Swype and Fleksy, were installed on the smartwatch. Swype is a QWERTY keyboard featuring auto-correction and a trace style of text input. Unlike traditional tapping, users enter text by connecting all letters in a word with one continuous trace. When the user lifts their finger from the trace, a word-selection and auto-correction algorithm is used to select the best matching word for the trace. A space is added at the end of the word once a new trace begins. Fleksy is a point-and-tap QWERTY style keyboard featuring auto-correction and onscreen

gestures to ease typing. A swipe-right gesture is used to input a space, swipe-left is used to delete, and swipe-up/down is used to scroll through suggested words.

Auditory and vibrotactile feedback were disabled on both keyboards. The order of the input methods was counterbalanced across participants. Forty randomly selected phrases from a list of 500 phrases composed by MacKenzie & Soukoreff [11] were used in the study. The phrases contained letters only (no digits, symbols, punctuation, or uppercase letters). Phrases ranged from 16 to 42 characters on both devices. Participants typed 20 phrases using each input method.

3.3   Procedure

After informed consent was obtained, participants completed a demographical survey and were trained on each input method. Participants then practiced test phrases with each input method for at least 5 min, and were allotted extra time if needed. Once they were comfortable with the input method, the experimental trials began. Phrases were displayed one at a time on a desktop computer. Participants read each phrase aloud to ensure comprehension, and verbally indicated when they started and stopped typing. Participants were asked to type as quickly and as accurately as possible and were not permitted to correct any errors. A researcher recorded task completion time using a stopwatch. Participants completed the System Usability Scale (SUS [12] and the NASA-TLX perceived workload assessment [13] after each keyboard condition. The order of input method condition was counterbalanced.

At the end, participants rated the two input methods on a 1–50 scale for preference, perceived speed and perceived accuracy. Finally, participants' hand dimensions were measured.

3.4   Dependent Variables

Typing Performance. Performance was measured by words per minute (WPM), adjusted words per minute (adjWPM), and word error rate (WER). adjWPM was calculated as WPM * WER. WER included substitution, insertion, and omission error rate.

Perceived Usability. The System Usability Scale (SUS) was used to measure participant's perceived usability of the keyboards. The SUS is an industry-standard 10-item scale with 5 response options (Strongly Disagree to Strongly Agree). The scale yields a score between 0–100, with higher score indicating higher usability.

Subjective Mental Workload. The raw NASA-TLX was used to measure participant's perceived workload of typing with each keyboard. Participants rated on a 20-point scale for mental, physical, temporal, performance, effort, and frustration.

Perceived Performance. Preference, perceived accuracy, and perceived speed with each keyboard were measured using a 50-point scale with 50 being most preferred.

4   Results

4.1   Typing Performance

A series of paired-sample t-tests were conducted to compare words per minute, adjusted words per minute and error rates across the two input methods. The first five trials were considered practice and eliminated from the analysis, leaving 15 test phrases for each method. Total error rate in addition to a breakdown by substitution, insertion, and omission errors was calculated for each keyboard at the word level. Cohen's d was used to measure effect size ($> .8$ = large) [14].

Typing speed measured by WPM revealed that participants typed faster using Swype than Fleksy, $t(17) = 5.54$, $p < .001$, Cohen's $d = 1.25$. The adjusted word per minute, which accounted for errors, produced similar results (Fig. 1). Total word error rate did not differ across keyboards, $t(17) = 1.86$, $p > .05$. The omission error rate, however, was smaller for Swype than for Fleksy, $t(17) = -2.94$, $p < .01$, Cohen's $d = 1.07$. Substitution and insertion error rates were not significantly different, $p > .05$.
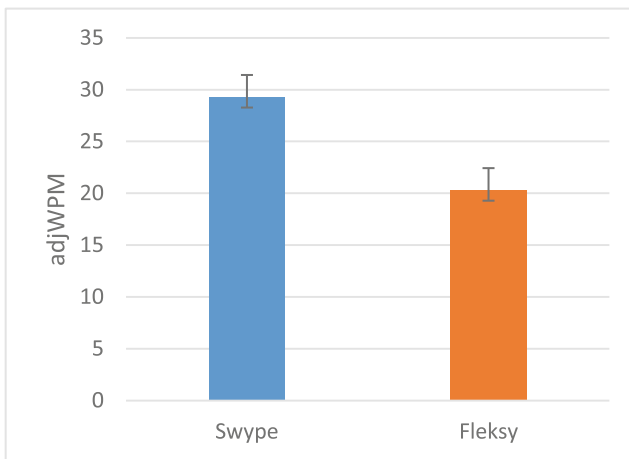

Fig. 1. Typing performance by keyboard

4.2   Subjective Measures

The perceived usability score measured by the System Usability Scale showed that participants rated Swype (M = 77.92, SD = 13.24) more usable than Fleksy (M = 46.61, SD = 17.47), t(17) = 6.21, p < .001. Fleksy was rated significantly more mentally t(17) = 3.79, p = .001 and physically demanding t(17) = 2.98, p = .008, more effortful t(17) = 4.29, p = .001 and frustrating t(17) = 5.55, p < .001 than Swype. Participants also reported that they performed worse with Fleksy than Swype, t(17) = 3.31, p = .004. Swype was perceived to be more accurate t(17) = 4.93, p < .001, faster t(17) = 5.37, p < .001, and was more preferable than Fleksy t(17) = 6.61, p < .001. The correlation between index finger width (M = 13.89 mm, SD = 1.20), which was used by

all participants, and adjWPM was not significant for either keyboard, r(17) = −0.409, p > .05 for Swype and, r(17) = −0.288, p > .05 for Fleksy, which suggested that typing performance would not vary across finger width significantly.

5   Discussion

This study demonstrated that manual text input on a smartwatch is feasible and in fact, can be quite efficient. Participants were able to type an average of 29.3 WPM with Swype and 20.3 WPM with Fleksy, after correcting for errors. Total error rates were approximately 9 % for the Swype and 16 % for the Fleksy, with the majority of errors resulting from incorrect word substitutions. The typing speed on a smartwatch is comparable to the typing speed of novice smartphone users. Castellucci et al. (2011) reported that the error rate of Swype on a smartphone was 7.0 % and the entry speed was 20.9 WPM [15].

The input performance with both Swype and Fleksy was faster than most reported performance using alternative keyboards on small screen devices. This indicates that the use of the familiar QWERTY-style input is a viable method of input, despite the fact that the target key sizes (4.10 mm for Swype and 5.66 mm for Fleksy) were smaller than recommended 7 mm key size for adequate performance [5].

It can be argued that the performance difference found in this study was mainly due to a superior algorithm used to autocorrect and suggest text alternatives with Swype. While the Fleksy key size was slightly wider and taller than the Swype key size, the entire keyboard size was comparable. It is not known how much of the superior performance of the Swype was due to the tracing method rather than the traditional tap that used with Fleksy. More research must be done to isolate performance differences due to the input technique.

References

1. Siegal, J.: Smartwatch sales set to explode, expected to top 100 M within four years. BGR (2013). http://bgr.com/2013/09/27/smartwatch-sales-forecast-2020/. 27 September 2013
2. Danova, T.: Smart watch forecast: The Apple Watch Mat Lead Wrist-worn Computers to Mass Market Status., Business Insider (2014). http://www.businessinsider.com/global-smartwatch-sales-set-to-explode-2014-3. 9 September 2014
3. Raghunath, M.T., Narayanaswami, C.: User interfaces for applications on a wrist watch. Pers. Ubiquit. Comput. **6**(1), 17–30 (2002)
4. Baudisch, P., Chu, G.: Back-of-device interaction allows creating very small touch devices. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 1923–1932. ACM (2009)
5. Funk, M., Sahami, A., Henze, N., Schmidt, A.: Using a touch-sensitive wristband for text entry on smart watches. In: CHI 2014 Extended Abstracts on Human Factors in Computing Systems, pp. 2305–2310. ACM (2014)
6. Oney, S., Harrison, C., Ogan, A., Wiese, J.: 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, pp. 2799–2802. ACM (2013)

7. Ashbrook, D., Lyons, K., Starner, T.: An investigation into round touchscreen wrist watch interaction. In: Proceedings of the 10th International Conference on Human Computer Interaction with Mobile Devices and Services, pp. 311–314. ACM (2008)
8. Xiao, R., Laput, G., Harrison, C.: Expanding the input expressivity of smartwatches with mechanical pan, twist, tilt and click. In: Proceedings of The 32nd Annual ACM Conference on Human Factors in Computing Systems, pp. 193–196. ACM (2014)
9. Bi, X., Smith, B.A., Zhai, S.: Quasi-qwerty soft keyboard optimization. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 283–286. ACM (2010)
10. Kim, S., Sohn, M., Pak, J., Lee, W.: One-key keyboard: a very small QWERTY keyboard supporting text entry for wearable computing. In: Proceedings of the 18th Australia Conference on Computer-Human Interaction: Design: Activities, Artefacts and Environments, pp. 305–308. ACM (2006)
11. MacKenzie, I.S., Soukoreff, R.W.: Phrase sets for evaluating text entry techniques. In: CHI 2003 Extended Abstracts on Human Factors in Computing Systems, pp. 754–755. ACM (2003)
12. Brooke, J.: SUS-A quick and dirty usability scale. In: Jordan, P., Thomas, B., Weerdmeester, B. (eds.) Usability Evaluation in Industry, pp. 189–194. Taylor & Francis, London (1996)
13. Hart, S.G., Staveland, L.E.: Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. In: Hancock, P.A., Meshkati, N. (eds.) Human Mental Workload. North Holland Press, Amsterdam (1988)
14. Cohen, J.: Statistical Power Analysis for Behavioral Sciences. Lawrence Erlbaum, New Jersey (1988)
15. Castellucci, S.J., MacKenzie, I.S.: Gathering text entry metrics on android devices. In: CHI 2011 Extended Abstracts on Human Factors in Computing Systems, pp. 1507–1512. ACM (2011)