In Only 3 Minutes: Perceived Exertion Limits of Smartwatch Use

Rushil Khurana
Carnegie Mellon University
Pittsburgh, PA
rushil@cmu.edu

Nikola Banovic
Carnegie Mellon University
Pittsburgh, PA
nbanovic@cs.cmu.edu

Kent Lyons
Tesla
Palo Alto, California
dr.kent.lyons@gmail.com

ABSTRACT
Glanceability and low access time are arguably the key assets of a smartwatch. Smartwatches are designed for, and excel at micro-interactions– simple tasks that only take seconds to complete. However, if a user desires to transition to a task requiring sustained usage, we show that there are additional factors that prevent possible longer usage of the smartwatch. In this paper, we conduct a study with 18 participants to empirically demonstrate that interacting with the smartwatch on the wrist leads to fatigue after only a few minutes. In our study, users performed three tasks in two different poses while using a smartwatch. We demonstrate that only after three minutes of use, the change in perceived exertion of the user was anchored as "somewhat strong" on the Borg CR10 survey scale. These results place an upper bound for smartwatch usage that needs to be considered in application and interaction design.

Author Keywords
smartwatches; perceived exertion; sustained use

ACM Classification Keywords
H.5.2 Information Interfaces and Presentation: User Interfaces; Evaluation/methodology

INTRODUCTION
Small wrist-worn devices such as smartwatches are becoming commonplace. Smartwatch interaction design is focused on micro-interactions where the users quickly attends to the watch, perform their tasks, and then disengage with the watch to resume other activities. Past research has shown the promise of interaction design focusing on micro-interactions because the position of the watch on the body affords very quick access times and glanceability [1]. Recent research has proposed improvements to smartwatch interactions that enable other, longer, and more involved kinds of tasks that users often perform on their mobile phones (e.g., text entry [9], [14]). Research also looks beyond current smartwatch hardware to expand the interaction area with the goal of enabling the users to perform more complex tasks [7].

ISWC '18, October 8–12, 2018, Singapore, Singapore

DOI: https://doi.org/10.1145/3267242.3267285

However the user has to hold their hand in a raised position [13] to use a smartwatch. For in-air gestures, holding the arm up for an extended period of time can cause fatigue in the upper limbs, referred to as the *gorilla-arm effect* [8]. Excessive fatigue leads to discomfort, impacts the ability of a user to complete a task, and may even pose health risks. Ultimately, if this effect is present for smartwatch use, it could reduce the usability of a wearable system because comfort is one of the most important considerations when designing for wearables [12] [10] [19]. The overall adoptability of a wearable device relies on both its functionality and perceived comfort [4]. And for extend duration of use, perceived comfort becomes an even more important consideration [16].

Understanding the users' perceived exertion while using a smartwatch helps make better design decision that improve overall interaction with the device. Although early work assessed comfort levels for head worn equipment such as helmets [15] and arm worn device [17], there is little research on exertion while using wearables. Knight and Baber determined that simply adding a wearable computer to the body leads to increase in perception of exertion and discomfort, and adopting the posture (disposition) to interact with it, leads to increased muscular activity [11]. Hincapié-Ramos *et al.*[8] developed a method to quantify arm fatigue in mid-air interactions (gestures) and termed it *consumed endurance*. While their work provides a framework for studying exertion, little is known about fatigue when interacting with a smartwatch.

In this paper, we investigate the impact of sustained usage of the watch to determine the change in perceived exertion with increase in usage time. We show that only after 3 (while standing) to 4 minutes (while sitting) of smartwatch use, a user exhibits *somewhat strong* exertion, making smartwatches in their current state, unsuitable for longer interactions.

Borg CR10 Scale for Perceived Exertion
To measure perceived exertion, we use the self-report Borg CR10 Scale [5]. It is a category (C) ratio (R) scale *i.e.*, twice a rating is likely to be twice as hard. It is a 12 point scale from 0 to 10 with an additional anchor at 0.5 with standard labels to ensure approximately equidistant scale intervals (Table 1). The scale has been studied, used, and verified in the literature to have a strong linear correlation with physiological indicators of perception such as heart rate [6], and thus a suitable measure for perceived exertion. It has previously been used in similar studies to measure fatigue in gestures [3] and discomfort of wearables [12].

Rating	Description
0	Nothing at all
0.5	Extremely Weak (Just noticeable)
1	Very Weak
2	Weak (Light)
3	Moderate
4	Somewhat Strong
5	Strong (heavy)
6	
7	Very Strong
8	
9	
10	Extremely Strong

Table 1. Borg CR10 Perceived Exertion Scale

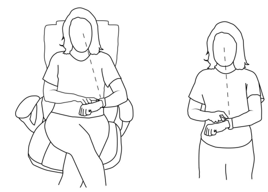

Figure 1. The two poses in which our study was conducted, (A) sitting with elbows resting on armrest, and (B) standing with arm raised.

Effect (main or interaction)	p-value
time	<0.001
pose	<0.001
input	<0.001
pose:input	<0.01
pose:time	<0.001
input:time	0.98
pose:input:time	0.68

Table 2. Table summarizing results of ANOVA with pose, input and time as independent variables and perceived exertion as the dependent variable after performing the ART procedure.

PERCEIVED EXERTION STUDY

Here, we investigate the impact of smartwatch use on arm fatigue. We hypothesize that using a smartwatch does not afford continuous use for long because users need to hold their arm up in order to see and touch the smartwatch.

Study Design

We conduct a within-subject study, which simulates typical users' target selection tasks with a smartwatch in different poses over time. We simulate it using three different forms of abstract smartwatch input primitives:

1. **Touch:** Participants tap a target once with their finger.
2. **Dwell:** Tap and hold a target for a specified time (500ms).
3. **Swipe:** Swipe a target in a particular direction (shown on screen) for a specified movement of at least 35 pixels.

The participants perform the inputs in two different poses (Figure 1), which mimicks commonly held smartwatch usage positions: A) sitting with elbows rested; and B) standing with arm raised. We denote each combination of input and pose as a condition in our study. The participants completed eight 30-second trials in each condition.

We measure participants' self-reported exertion with the Borg CR10 scale [5], once before each condition to establish a baseline, and at the end of each 30 second trial. Participant provided their responses verbally and a researcher recorded their response on paper.

We recruited 18 participants (8 female, 10 male) via word of mouth. The participants' ages ranged from 22 to 35 (median = 25.5) and their smartwatch experience varied from non-users to power users.

Procedure

Participants start the study in one of two poses (sitting or standing). The poses are counter balanced across participants and the conditions (input task) within each pose are randomized. Participants complete 8 trials of 30 seconds in a given condition with a 5-minute break after each condition to rest and to reduce any carryover effects.

For each trial, participants select a target as many times as they can in the allotted time. Touch and dwell targets are solid circles with a 35 pixel radius, while swipe targets have arrows indicating a particular direction participants swipe in to select the target. If the participants accurately select the target, it disappears and the next target appears at a random position on the screen. We time-stamp and log all touch events, the count of accurate touches, and total touches during each trial.

We used the Samsung Gear Live smartwatch with a screen size of 1.63 inches, and a resolution of 320x320 pixels. The watch runs Android Wear 6.0. We developed a custom Android application that progresses through our experimental procedure. The application pauses every 30 seconds to allow researchers to collect participant self-reported exertion and any other feedback.

RESULTS

We report our study results investigating impact of different poses, input primitives, and usage time on user exertion.

We collect our data using the ordinal Borg CR10 Scale. The data was not normally distributed according to a Shapiro-Wilk test ($p<0.05$). We transform our data using ART analysis and the associated ARTool [18]. This transformation aligns the perceived exertion scores for each main or interaction effect and assigns them ranks. This procedure allows us to conduct a parametric three-way repeated measures ANOVA (INPUT x POSE x TIME) for each effect on the transformed data [18]. Table 2 shows the results of our ANOVA. We present results of additional post-hoc tests on all significant effects below.

Usage Time- There is a significant effect of usage time on exertion. The raw exertion scores of participants show a consistent upward trend over time. Figure 2 and 3 show the boxplots of the raw (non-transformed) exertion values in each input primitive in both poses. Thus, we can establish that an increase in usage time leads to increase in perceived exertion.

On the transformed data, we conduct a post-hoc Tukey's pairwise comparisons at each time point and observe a significant difference ($p<0.05$) between all pairs of time points except between 3.5 and 4 minutes. Moreover, the least-squared mean value of exertion increases with increase in time. Together, it confirms that participants consistently reported higher perceived exertion with increase in usage time.

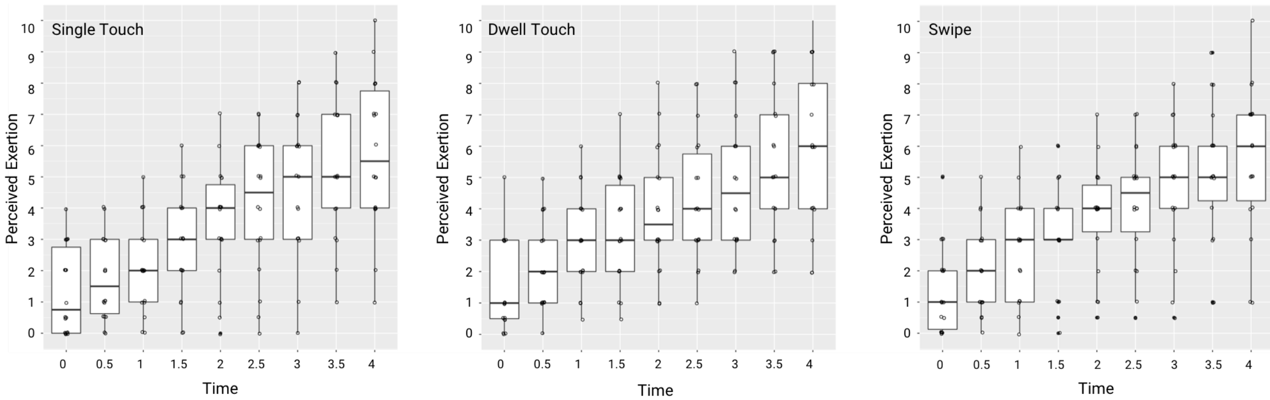

Figure 2. Boxplot and linear regression model for PE of subjects in standing pose for all three conditions: Single (left), Dwell (middle), and Swipe (right)

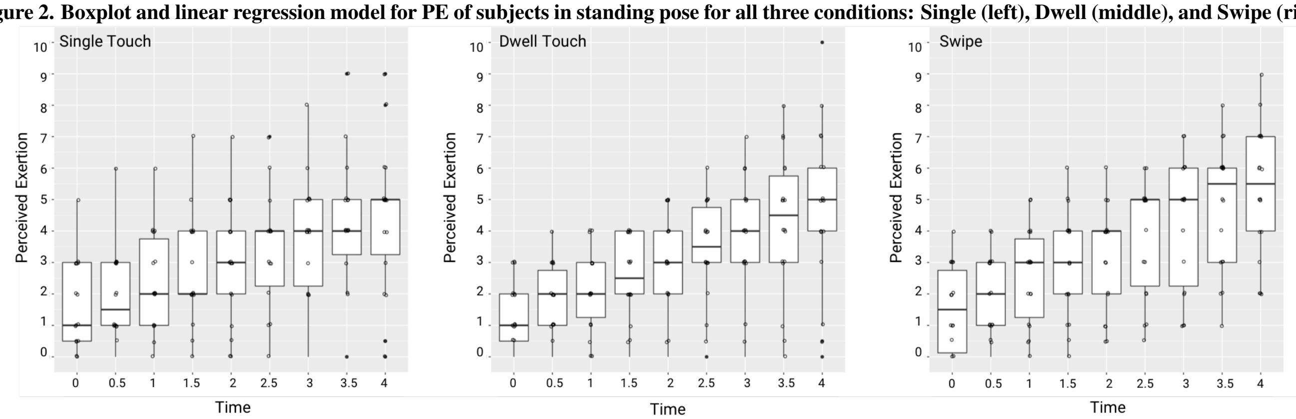

Figure 3. Boxplot and linear regression model for PE of subjects in sitting pose for all three conditions: Single (left), Dwell (middle), and Swipe (right)

We also want to ascertain after how much time does a smartwatch use becomes unsuitable for continuous interactions. A self-reported value of 4 on the CR10 scale is anchored as 'somewhat strong'. We use this as a threshold for when continued usage would become a significant usability issue.

In the standing pose, the median exertion score exceeds 'somewhat strong' at t=2.5min for Single Touch and Swipe. It exceeds 'somewhat strong' at t=3min for dwell. In the sitting pose, it exceeds this threshold at t=4min for Single Touch, t=3.5min for Dwell, and t=2.5min for Swipe.

For any form of input primitive, a user is expected to be 'somewhat strongly' tired after only 3 minutes of smartwatch use in the standing pose, and 4 minutes of use when sitting.

Pose- A significant effect of pose on exertion indicates that the kind of pose held by the participants makes a difference in their perceived exertion. We conduct a post-hoc Tukey's pairwise comparison of the two poses within our pose variable: standing versus sitting. A significant difference ($p<0.0001$) exists between the two poses. The standing pose has a higher least-squared mean exertion, which means participants were exerted more using the smartwatch when standing.

Pose:Time- The relationship between pose and time is important to better understand the individual effects of both factors. So, we fix pose and look at the main effects of time. For both poses, we obtain a significant effect of time ($p<0.001$) with ANOVA. It means, even though there are differences in how much time affects exertion in each pose, regardless, time has a significant effect on the perceived exertion.

For both poses, we also conduct a post-hoc Tukey's pairwise comparisons at each time point. In the sitting pose, a 30 second difference in time pairs did not reveal significance for any pair except between 0.5 minutes and 1 minute. However, a difference of 1 minute or greater led to significant differences in the exertion values between any two time points.

In contrast in the standing pose, all time pairs up until 2 minutes with a difference of 30 seconds revealed significance. After that, there was no observed significance between pairs that are 30 seconds apart. Again, a difference of 1 minute or greater led to significant differences in the exertion levels between any two time points.

What this means is that using the smartwatch while standing can lead to a massive increase in exertion quickly (~30 seconds) whereas it takes a little longer (~1 minute) for a noticeable change in exertion while sitting. It supports the results we saw earlier with the standing pose demanding higher physical effort than the sitting pose.

Input- A significant effect of the kind of input primitive on exertion indicates that the exertion varies by type of input. We conduct a post-hoc Tukey's pairwise comparison of the three primitives: Touch, Dwell, and Swipe. A significant difference exists between Touch and Swipe ($p<0.05$), and Dwell and Swipe ($p<0.001$). Swipe has the highest least-squared mean exertion value and Single Touch has the lowest least-squared mean exertion value across all participants. This indicates that the Swipe primitive, commonly used on touch devices takes the most perceived exertion to use on a smartwatch.

Pose:Input- There is a significant interaction effect between pose and input, but we did not explore it further. The relationship between them is interesting and warrants further exploration, however is not pertinent to our hypothesis.

Other Measures

Our analysis results are also corroborated with observations from the study and subjective feedback by the participants. Some participants were noticeably tired, and some performed arm stretches/exercises in the rest-time between conditions citing fatigue in their arm muscle. As potential side effects of exertion, we looked at accuracy and average reaction time as metrics that can be affected by time. While the accuracy shows variance depending on the input primitive, we observe an overall decrease in the average reaction time for the dwell input primitive in both poses. However, there was no significance found in time as a predictor for accuracy, or average reaction time in any of the cases.

DISCUSSION & CONCLUSION

Our study provides evidence that smartwatches are not suited for extended interactions in their current form. The manufactures provide hints to placing intentional design limitations on current smartwatch applications and their user interfaces. Android Wear guidelines state that an app should be designed for glanceability. Their guidelines are intentionally targeted at applications that only require short usage.

The QuickDraw study [1] found the lower bounds of watch interactions based on the time required to access the device. Our study places an upper bound between 3 and 4 minutes before perceived endurance starts to impact usage.

As we explore ways to overcome the technical and interaction limitations of current smartwatches to enable richer and fuller application experiences, we also need to consider the results of this study. Banovic *et al.* found that a large portion of the mobile interactions happen in 5 minutes [2]. Even if some of the other interaction challenges for smartwatches are addressed with additional research (*e.g.,* text entry speed), a large number of traditional phone interactions would still take too long on a smartwatch and lead to fatigue. If we want to enable longer smartwatch interactions, we need alternative techniques that do not require the user to hold and sustain the pose currently employed by smartwatches.

REFERENCES

1. Daniel L Ashbrook, James R Clawson, Kent Lyons, Thad E Starner, and Nirmal Patel. 2008. Quickdraw: the impact of mobility and on-body placement on device access time. In *Proc. of CHI*.
2. Nikola Banovic, Christina Brant, Jennifer Mankoff, and Anind Dey. 2014. ProactiveTasks: the short of mobile device use sessions. In *Proc. of MobileHCI*.
3. Kayne Barclay, Danny Wei, Christof Lutteroth, and Robert Sheehan. 2011. A quantitative quality model for gesture based user interfaces. In *Proc. of Australian CHI*.
4. Kerry Bodine and Francine Gemperle. 2003. Effects of functionality on perceived comfort of wearables. In *Proc. of ISWC*.
5. Gunnar Borg. 1998. *Borg's perceived exertion and pain scales.* Human kinetics.
6. Aaron J Coutts, Ermanno Rampinini, Samuele M Marcora, Carlo Castagna, and Franco M Impellizzeri. 2009. Heart rate and blood lactate correlates of perceived exertion during small-sided soccer games. *Journal of Science and Medicine in Sport* 12, 1 (2009).
7. Rajkumar Darbar, Prasanta Kr Sen, and Debasis Samanta. 2016. PressTact: Side Pressure-Based Input for Smartwatch Interaction. In *Proc. of the CHI Extended Abstracts*.
8. Juan David Hincapié-Ramos, Xiang Guo, Paymahn Moghadasian, and Pourang Irani. 2014. Consumed endurance: a metric to quantify arm fatigue of mid-air interactions. In *Proc. of CHI*.
9. Jonggi Hong, Seongkook Heo, Poika Isokoski, and Geehyuk Lee. 2015. SplitBoard: A simple split soft keyboard for wristwatch-sized touch screens. In *Proc. of CHI*.
10. Richard Hurford, Adam Martin, and Povl Larsen. 2006. Designing wearables. In *Proc. of ISWC*.
11. James F Knight and Chris Baber. 2000. Wearable computers and the possible development of musculoskeletal disorders. In *Proc. of ISWC*.
12. James F Knight, Daniel Deen-Williams, Theodoros N Arvanitis, Chris Baber, Sofoklis Sotiriou, Stamatina Anastopoulou, and Michael Gargalakos. 2006. Assessing the wearability of wearable computers. In *Proc. of ISWC*.
13. Kent Lyons and Halley Profita. 2014. The multiple dispositions of on-body and wearable devices. *Pervasive Computing* 13, 4 (2014).
14. Franck Poirier and Mohammed Belatar. 2015. UniWatch-some approaches derived from UniGlyph to allow text input on tiny devices such as connected watches. In *International Conference on Human-Computer Interaction*.
15. Kathleen M Robinette and Jennifer J Whitestone. 1994. The need for improved anthropometric methods for the development of helmet systems. *Aviation, space, and environmental medicine* 65, 5 Suppl (1994).
16. Thad Starner. 2014. How wearables worked their way into the mainstream. *Pervasive Computing* 13, 4 (2014).
17. Robert Stein, Stephen Ferrero, Margaret Hetfield, Alan Quinn, and Mark Krichever. 1998. Development of a commercially successful wearable data collection system. In *Proc. of ISWC*.
18. 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 *Proc. of CHI*.
19. Joanne E Zucco and Bruce H Thomas. 2016. Design guidelines for wearable pointing devices. *Frontiers in ICT* 3 (2016).