# Evaluating Tap, Flick, Force, and Wrist Gestures for Interacting with Vertical Displays via Smartwatch

Yuan Ren
Inclusive Interaction Lab
University of California, Merced
Merced, CA, USA
yren5@ucmerced.edu

Ahmed Sabbir Arif
Inclusive Interaction Lab
University of California, Merced
Merced, CA, USA
asarif@ucmerced.edu

## Abstract

This work compares four methods for interacting with large vertical displays via a smartwatch. In addition to conventional taps, it includes directional flicks, wrist gestures, and force-based methods, which can reduce the need for precise target selection on small screens. Results show that directional flicks performed nearly as well as taps in terms of speed, error rate, and perceived usability. Surprisingly, force-based input outperformed wrist gestures in both usability and perceived task load, despite concerns about users' ability to control pressure. Wrist gestures, though slower and more demanding, were preferred by some users for one-handed use.

## Keywords

Indirect Interaction, Gestural Interaction, Force, Television

**ACM Reference Format:**
Yuan Ren and Ahmed Sabbir Arif. 2025. Evaluating Tap, Flick, Force, and Wrist Gestures for Interacting with Vertical Displays via Smartwatch. In *The PErvasive Technologies Related to Assistive Environments (PETRA '25), June 25–27, 2025, Corfu Island, Greece.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3733155.3736602

## 1 Introduction

Interacting with large vertical displays can be challenging due to their size. Users often cannot physically reach all areas of the display, making interaction inconvenient or even impossible. Prior research has explored alternative methods using external devices such as smartphones, tablets, laser pointers, or gesture-based systems [1]. However, these approaches are often inaccessible to individuals with motor or other physical disabilities. Smartwatches offer a promising alternative. Worn on the wrist, they are always within reach and do not require users to perform large or expressive gestures. However, their small screens and reliance on tap-based interaction make them difficult to use for precise target selection, especially for users with motor impairments [6]. In this work, we investigate alternative input methods for smartwatch-based interaction with large vertical displays. Our goal is to identify techniques that reduce the need for precise targeting and can be performed without using the other hand or fingers.



*PETRA '25, Corfu Island, Greece*

ACM ISBN 979-8-4007-1402-3/25/06
https://doi.org/10.1145/3733155.3736602

## 2 Comparative Study

We conducted a user study to compare the following input methods. For each task, we determined the interaction thresholds through a pilot study conducted with three participants (1 female, 2 male, all right-handed, M = 30 years, SD = 1), using the experimental device.

**Tap:** This method included two tasks: regular tap and long tap (also known as long press or press & hold). For long tap, we set a threshold of 2 seconds, meaning participants had to maintain touch for at least two seconds to activate it. We did not include double-tap, as prior studies have shown that users with motor disabilities often find it difficult [6].

**Directional Flick:** In this method, users performed flick gestures (up, down, left, or right) anywhere on the screen. A movement threshold of 10 pixels was used, requiring the finger to travel at least this distance to register a flick.

**Wrist Gesture:** Here, users twisted their wrist toward their body (down), as if checking the time, or away from their body (up). We defined a flat position within the range of 0.7 to 10.97 degrees. Any angular deviation beyond this range in either direction was registered as a twist.

**Force-Based:** In this method, users applied different levels of pressure: soft, regular, or hard, when tapping. Taps with a force below 0.01 and a duration over 132 ms were considered as soft, and those with a force above 0.02 and a duration over 551 ms were considered as hard.

**Figure 1: A participant interacting with a vertical display.**

### 2.1 Participants

Twelve participants took part in the study (M = 26.83 years, SD = 6.1). Seven identified as female and five as male. All participants were right-handed and wore the smartwatch on their left wrist. Eight were smartwatch owners (M = 3 years, SD = 2.4), while the remaining four, though not owners, were familiar with smartwatch use. Each participant received US $10 for their participation.

## 2.2 Apparatus

We used a 1,651 mm LG UL3G-B Series commercial display monitor with a resolution of 3,840 × 2,160 (UHD), running on webOS 4.0 (Fig. 1). For the smartwatch, we used the Fossil Gen 6 FTW4061V (44 mm, 118 g), which features a 416 × 416 resolution at 326 ppi. It is powered by a Qualcomm Snapdragon Wear 4100+ processor, with 8 GB of storage, 1 GB of RAM, and built-in Bluetooth 5.0 LE for seamless compatibility with the vertical display.

## 2.3 Design

We used a within-subjects design for the study. The independent variable was method, which had four levels, each with multiple sub-levels (tasks): tap (regular tap, long tap), directional flick (up, down, left, right), wrist gesture (up, down), and force-based (soft, regular, hard). Each participant performed 20 trials per sub-level, resulting in a total of 12 participants × 11 tasks × 20 trials = 2,640 interactions. The dependent variables were the following performance metrics:

**Task completion time**: The average time (in seconds) participants took to accurately complete each task (sub-level).

**Error rate**: The average number of errors made per task. An error was recorded whenever the participant's input did not match the target sub-level.

## 2.4 Procedure

We began by introducing the study to participants, obtaining informed consent, and collecting demographic information. Participants were then given time to practice each input method by performing each task twice while seated. The main study followed immediately after. The vertical display, mounted on a stand 90 cm above the ground and positioned 360 cm from the participant, presented one task at a time in random order. Participants were instructed to perform each task without looking at the smartwatch. Upon successful completion, the display provided visual feedback by changing the font color from white to green and then displayed the next task. Error correction was enforced, that is, participants had to repeat each task until it was accurately recognized by the system. All interactions were logged automatically. After completing the study, participants filled out a custom usability questionnaire and the NASA-TLX workload assessment [4].

# 3 Results & Discussion

We conducted repeated-measures ANOVAs for all analyses of quantitative data and reported effect sizes ($\eta^2$) for all statistically significant results. For non-parametric questionnaire data, we used the Friedman test, along with Kendall's coefficient of concordance ($W$) to report effect sizes. We also filtered the data by method to examine the effect of different tasks on performance. For methods with two tasks (sub-levels), we conducted t-tests and reported Cohen's $d$ as the effect size.

## 3.1 Task Completion Time

An ANOVA revealed a significant effect of method on task completion time ($F_{3,11} = 11.28, p < .0001, \eta^2 = 0.04$). Post hoc Tukey-Kramer analysis identified two distinct groups: {Directional Flick} and {Tap, Wrist Gesture, Force-Based}, with the directional flick

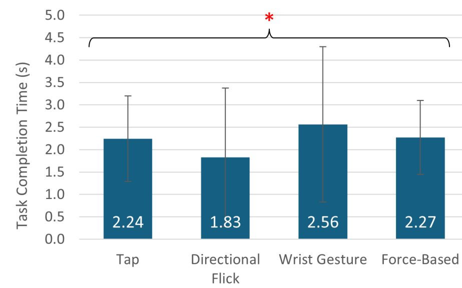



**Figure 2: Average task completion time with ±1 standard deviation. Red asterisks indicate significant differences.**

method being significantly faster than the others. Fig. 2 illustrates the average task completion time across the four methods.

*3.1.1 Tap.* A paired samples t-test revealed a significant effect of task on task completion time ($t = 21.08, df = 239, p < .0001, d = 1.36$). Long tap took 78% longer to perform than regular tap (Fig. 3a), which is expected given its inherently slower design.

*3.1.2 Directional Flick.* An ANOVA found no significant effect of flick direction on task completion time ($F_{3,11} = 0.05, p = .98$), suggesting that flick direction had no impact on performance speed. In fact, flicks in all directions took approximately 1.8 seconds to complete (Fig. 3b).

*3.1.3 Wrist Gesture.* A paired samples t-test found no significant effect of twist direction on task completion time ($t = -0.81, df = 239, p = 0.42$). Participants took approximately 2.5 seconds to perform wrist twists in both the up and down directions (Fig. 3c).

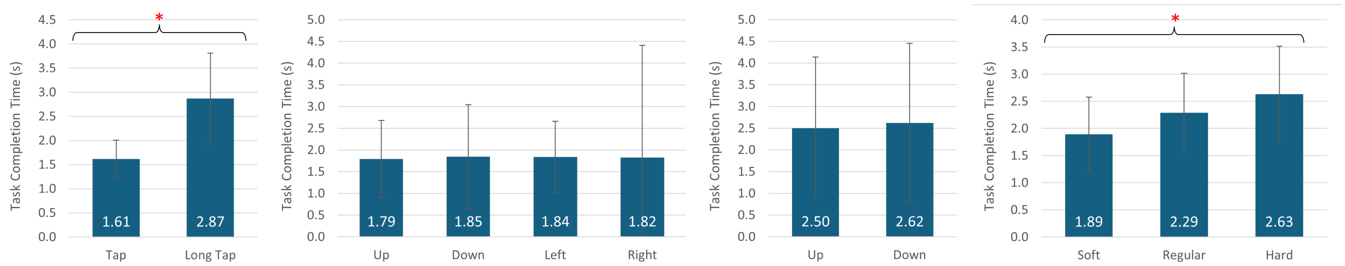
*3.1.4 Force-Based.* An ANOVA revealed a significant effect of force level on task completion time ($F_{2,11} = 26.04, p < .0001, \eta^2 = 0.13$). Post hoc Tukey-Kramer analysis identified two distinct groups: {Soft, Regular} and {Hard}, with the hard press taking significantly longer to perform than the other levels. Participants were 28% faster with soft force and 13% faster with regular force compared to hard force (Fig. 3d).

## 3.2 Error Rate

An ANOVA also revealed a significant effect of input method on error rate ($F_{3,11} = 28.93, p < .0001, \eta^2 = 0.08$). The Tukey-Kramer test identified two distinct groups: {Force-Based} and {Tap, Directional Flick, Wrist Gesture}, with the force-based method resulting in significantly higher error rates than the other methods. Fig. 4 shows the average error rate across all four methods.

*3.2.1 Tap.* Participants made no errors when performing tap and long tap (Fig. 5a), most likely due to their familiarity with these actions from regular use on various touchscreen-based devices.

*3.2.2 Directional Flick.* An ANOVA revealed a significant effect of direction on the error rate of directional flicks ($F_{3,11} = 3.06, p < .05, \eta^2 = 0.02$). Post hoc Duncan's test identified three distinct groups: {Down, Up}, {Right}, and {Left}, with Down and Up being the least error-prone, and Left being the most error-prone (Fig. 5b).




**(a) Tap** **(b) Directional Flick** **(c) Wrist Gesture** **(d) Force-Based**

**Figure 3: Average task completion time for each task within each method. Red asterisks indicate statistically significant differences. Error bars represent ±1 standard deviation.**

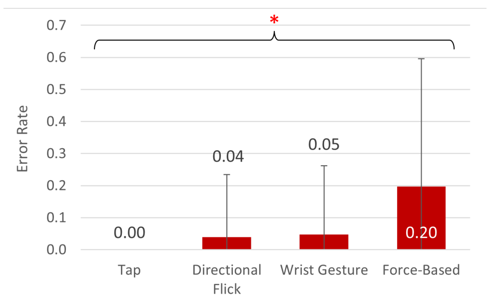



**Figure 4: Average error rate with ±1 standard deviation. Red asterisks indicate significant differences.**

This may be attributed to the fact that all participants were right-handed. Interestingly, however, these errors did not impact the overall task completion time for the Left direction.

*3.2.3 Wrist Gesture.* A paired samples t-test revealed a significant effect of twist direction on error rate ($t = 3.52, df = 239, p < 0.001, d = 0.23$). Although overall error rates were low (below 1.0), participants made 75% more errors when twisting the wrist away from the body compared to toward the body (Fig. 5c). This is likely due to wrist physiology, which makes inward twisting easier. In addition, wristwatch users commonly perform this inward motion to check the time, contributing to greater familiarity.

*3.2.4 Force-Based.* An ANOVA revealed a significant effect of force level on error rate ($F_{2,11} = 3.69, p < .05, \eta^2 = 0.06$). Post hoc Tukey-Kramer analysis identified three distinct groups: {Soft}, {Regular}, and {Hard}, with Soft and Hard being significantly different from each other. Surprisingly, the Regular force level resulted in a much higher error rate−55% higher than Soft and 67% higher than Hard (Fig. 5d). Post-study comments suggest this may be due to having difficulty maintaining a consistent force at the Regular level.

## 3.3 Usability

A Friedman test identified a significant effect of method on perceived speed ($\chi^2 = 8.41, p < .05, W = 0.23$), perceived accuracy ($\chi^2 = 22.43, p < .0001, W = 0.62$), intuitiveness ($\chi^2 = 8.66, p < .05, W = 0.24$), naturalness ($\chi^2 = 15.91, p < .001, W = 0.44$), ease of use ($\chi^2 = 24.09, p < .0001, W = 0.67$), learnability ($\chi^2 = 12.18, p < .01, W = 0.34$), and willingness to use ($\chi^2 = 21.32, p < .0001, W = 0.59$). Since all results yielded medium ($W = 0.3$) to large ($W = 0.5$) [2], it is likely that the findings will generalize to a larger sample. Fig. 6 presents the average user ratings across all aspects.

Results indicate that participants found directional flicks comparable to taps across all evaluated aspects. In contrast, wrist gestures received the lowest ratings overall. Interestingly and somewhat unexpectedly, the force-based method was perceived moderately well, showing comparable ratings for speed, intuitiveness, and learnability compared to taps and directional flicks.

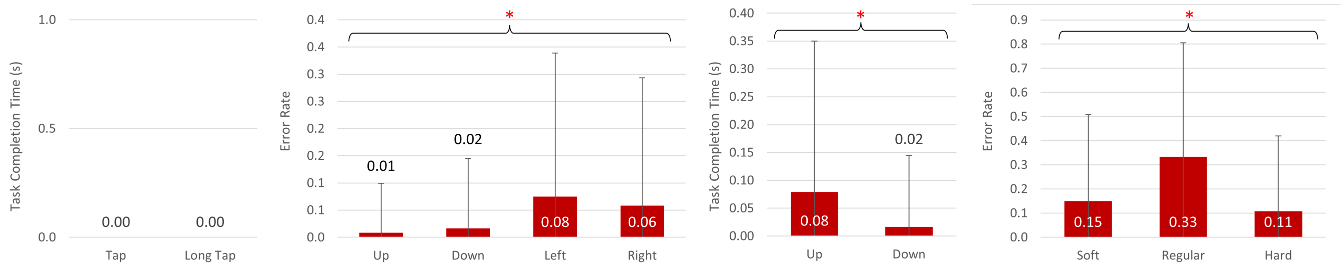
## 3.4 Perceived Workload

A Friedman test identified a significant effect of method on mental demand ($\chi^2 = 15.32, p < .005, W = 0.42$), physical demand ($\chi^2 = 15.50, p < .005, W = 0.43$), temporal demand ($\chi^2 = 18.13, p < .0005, W = 0.50$), performance ($\chi^2 = 15.39, p < .005, W = 0.43$), effort ($\chi^2 = 22.25, p < .0001, W = 0.62$), and frustration ($\chi^2 = 15.87, p < .005, W = 0.44$). Similar to the usability, all workload ratings yielded medium to large Kendall's $W$, suggesting that the findings are likely to generalize to a larger sample. Fig. 7 presents average user ratings for all workload aspects.

As anticipated, tap was rated as the most effortless and least frustrating, closely followed by directional flicks. In contrast, participants found the wrist gesture to be the most demanding, lowest in performance, and most frustrating, followed by the force-based method. Although wrist gestures were generally rated poorly in terms of perceived workload, some participants ($N = 4$) rated them comparably to tap. One participant (female, 42 years) even preferred it over the other methods, stating, *"The wrist gesture is the most desirable one, since it doesn't need the other hand [to] tap the watch."*

# 4 Implications & Conclusion

The results highlight directional flicks as a strong alternative to traditional tap-based interactions on smartwatches. Task completion times for directional flicks were nearly equivalent to taps, both averaging between 1.5 and 2 seconds. Flicks also resulted in the lowest error rate after taps, and participants rated them as nearly as usable and effortless. Unlike tap-based interactions, which require precise target selection (e.g., tapping buttons or icons), directional flicks can be performed anywhere on the screen. This makes them especially promising for improving accessibility for individuals with motor




**(a) Tap** **(b) Directional Flick** **(c) Wrist Gesture** **(d) Force-Based**

**Figure 5: Average error rate for each task within each method. Red asterisks indicate statistically significant differences. Error bars represent ±1 standard deviation.**

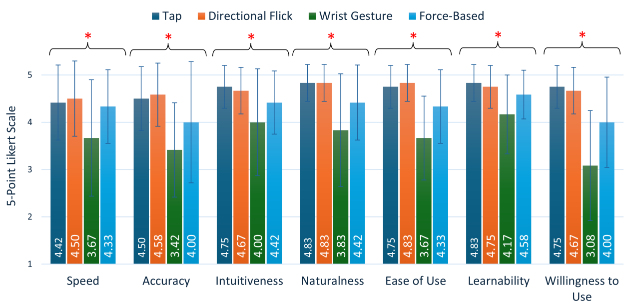



**Figure 6: Average user ratings on a 5-point scale. Error bars show ±1 SD. Red asterisks indicate statistical significance.**

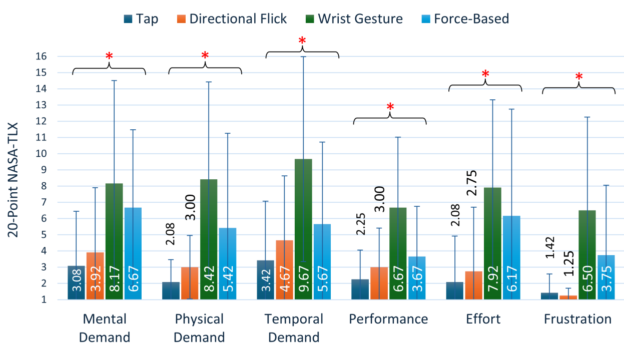



**Figure 7: Average Raw NASA-TLX ratings A 20-point scale. Error bars show ±1 SD. Red asterisks indicate indicate statistical significance.**

impairments. Supporting this, prior research has demonstrated the success of gesture-based input on larger devices like tablets [3, 5].

Interestingly, the force-based method performed better than anticipated. It was rated higher in usability and lower in perceived task load than wrist gestures. Although we expected it to perform poorly due to the challenge of applying consistent pressure, participants quickly adapted to the force thresholds. This is encouraging, as prior studies have shown that force input can also reduce the need for precise targeting by allowing users to select very small targets through force variation [7].

In contrast, the wrist gesture method did not perform as well as expected. It had the slowest task completion time, lower usability ratings, and higher perceived workload. These outcomes are likely influenced by the physical limitations of wrist movement. However, it was more accurate than the force-based method, and participants did not express disinterest in using it. In fact, some noted an initial learning curve but reported improvement over time. One participant (male, 24 years) commented, *"It was a little difficult at first, but I got used to it."* Another participant (female, 42 years) found wrist gestures to be the most desirable method, as they could be performed using just one hand. This one-handed interaction model may be especially valuable for users with motor or situational impairments [8], such as when the other hand is occupied.

These findings point to promising alternatives to tap-based interactions on smartwatches for improving accessibility and supporting more flexible, one-handed use.

## References

[1] Carmelo Ardito, Paolo Buono, Maria Francesca Costabile, and Giuseppe Desolda. 2015. Interaction with Large Displays: A Survey. *ACM Comput. Surv.* 47, 3 (Feb. 2015), 46:1–46:38. https://doi.org/10.1145/2682623

[2] Ahmed Sabbir Arif. 2017. *A Brief Note on Selecting and Reporting the Right Statistical Test.* Technical Report. University of California, Merced, United States. https://www.theiilab.com/notes/HypothesisTesting.html

[3] Ahmed Sabbir Arif, Michel Pahud, Ken Hinckley, and Bill Buxton. 2014. Experimental Study of Stroke Shortcuts for a Touchscreen Keyboard with Gesture-Redundant Keys Removed. In *Proceedings of Graphics Interface 2014 (GI '14)*. Canadian Information Processing Society, Toronto, ON, Canada, 43–50. http://dl.acm.org/citation.cfm?id=2619648.2619657

[4] Sandra G. Hart and Lowell E. Staveland. 1988. Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research. In *Advances in Psychology*. Vol. 52. Elsevier, 139–183. https://doi.org/10.1016/S0166-4115(08)62386-9

[5] Tomer Moscovich. 2009. Contact Area Interaction with Sliding Widgets. In *Proceedings of the 22nd annual ACM symposium on User interface software and technology (UIST '09)*. Association for Computing Machinery, New York, NY, USA, 13–22. https://doi.org/10.1145/1622176.1622181

[6] Gulnar Rakhmetulla and Ahmed Sabbir Arif. 2023. Crownboard: A One-Finger Crown-Based Smartwatch Keyboard for Users with Limited Dexterity. In *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems (CHI '23)*. Association for Computing Machinery, New York, NY, USA, 1–22. https://doi.org/10.1145/3544548.3580770

[7] Yuan Ren and Ahmed Sabbir Arif. 2023. Investigating a Force-Based Selection Method for Smartwatches in a 1D Fitts' Law Study and Two New Character-Level Keyboards. In *Proceedings of the Seventeenth International Conference on Tangible, Embedded, and Embodied Interaction*. ACM, Warsaw Poland, 1–10. https://doi.org/10.1145/3569009.3572741

[8] Jacob O. Wobbrock. 2019. Situationally Aware Mobile Devices for Overcoming Situational Impairments. In *Proceedings of the ACM SIGCHI Symposium on Engineering Interactive Computing Systems (EICS '19)*. Association for Computing Machinery, New York, NY, USA, 1–18. https://doi.org/10.1145/3319499.3330292