Is Bigger Better? A Fitts' Law Study on the Impact of Display Size on Touch Performance

Corinna List and Michael Kipp

Augsburg University of Applied Sciences, Germany
{corinna.list,michael.kipp}@hs-augsburg.de

Abstract. Touch-sensitive surfaces are already a standard form of interaction. These surfaces come in many different sizes like tablets or touch walls. However, there is little research to characterize the impact of surface size on touch performance. We conducted a Fitts' Law study of three display sizes (13.5" tablet, 28" monitor, 69.5" large monitor), comparing various performance measures. We found that the smallest size (13.5") is problematic both objectively (high error rate) and subjectively (imprecise, difficult-to-use). In contrast, both the medium (28") and the large display (69.5") perform equally well. However, small displays allow for greater interaction speed compared to very large screens. Our results can help interaction designers and automatic algorithms to optimize current and future touch devices.

Keywords: Touchscreen surfaces · Fitts' Law · Gestural Interaction.

1 Introduction

Interactive surfaces can be found almost anywhere and in any size. Due to advances in display technologies, screens are becoming both smaller and larger than ever before. Surfaces can be as small as a watch face or as big as a shop window. While it is clear that different sizes afford different interaction methods and approaches, it is less clear how precisely touch performance differs on various surface sizes. We focus on surfaces with a minimal size of a tablet where one would expect a touch movement performance that is relatively unaffected by the boundaries of the screen.

There is little research that focuses characterizing the different sizes with precise measurements. It has been shown that a larger display size negatively affects the performance when using a mouse [16]: With increasing display size, movement time increases and performance decreases. However, in another study [5] users preferred a large display for daily work because it helped them with multi-window or rich information tasks.

When using touch, does the above apply as well, and does performance decrease with increased display size? One study [3] compared different screen sizes for performance and ergonomics using specific postures for each display (sitting with the tablet, standing in front of the public display). This makes it difficult to interpret the results. Hence, we focus on making a performance comparison

with equal side conditions. For each condition (tablet, monitor, wall) the user would stand and operate on a vertical surface.

In this paper we focus on the characteristics of display sizes when using touch for basic translation movements. In a Fitts' Law [6] experiment with 16 Participants we compared three different display sizes (small 13.5", medium 28", large 69.5") regarding various performance measures such as error rate or speed. To allow comparison with other existing research, we follow Soukoreff and MacKenzie [13]: We use the Shannon formulation to calculate the Index of Difficulty (ID), we use ID values between 2 to 6. We analyzed error rate, speed, and performance. We also assessed subjective difficulty and precision for each display using questionnaires.

Our results show that a small screen (13.5") performs worst. It is not only perceived as difficult-to-use and imprecise but also has the highest error rate. Both the medium sized screen (28") and the large display (69.5") performed equally well. However, the small screen outperformed the largest screen in terms of speed.

First, we discuss various studies on the effect of display size on performance (Section 2). In Section 3, our user study is described with respect to participants, apparatus, study procedure, and measurements. Afterwards we show (Section 4) and discuss (Section 5) the results for each display and in comparison with each other. Section 6 presents our conclusion and future work.

2   Related Work

There is much research on the broader topic of touch performance. Research has looked at various touch techniques as well as side conditions like screen angle and the grip used with mobile devices. Touch techniques for translation, rotation and scaling have been investigated to understand the effect of parameters like direction and position [7, 8, 11, 12]. Mobile devices and especially the tap gesture have been another focus of attention [9].

Nguyen and Kipp [11, 12] investigated the efficiency of translation and rotation touch gestures. They compared directions and screen regions as well as the effect of the display's orientation. They found patterns of high performance, for instance that upward movements performed faster than downwards movements on a horizontal display. They also found that movement which employs a shorter kinematic chain is faster.

The efficiency of rotation as well as pinch touch gestures was investigated by Hoggan et. al. [7, 8]. They found that with increasing diameter the duration and failure rate increase as well. An increased distance led to increased duration and failure rate.

Although these studies investigated how different parameters of touch input affect performance, they did not include display size as a factor.

For smartphones there are a number of studies looking at the performance for different grips. Lehmann and Kipp [9] investigated the performance, precision, and errors of tapping gestures on smartphones. They showed that a two-handed

use while tapping with the thumbs in landscape orientation leads to the best performance. They did not compare other display sizes.

There is also a range of studies looking at screen size as one side condition [2, 3, 5, 14, 16]. However, none of the following studies exclusively focus on size.

Wang et al. [16] investigated the impact of display size on mouse performance in a pointing task. They used four sizes: 10", 27", 46", and 55". In a Fitts' Law experiment they found that display size has no significant effect on accuracy. However, the larger the display, the larger the necessary movement time and the higher the slope and intercept.

Anslow and Wong [2] examined the effects of display angle and physical size on large touch displays in the work place for data analysis activities. The 27" display was rated more effective than the 40" screen. Participants claimed the 40" screen to be too difficult for reaching and interacting effectively with all screen locations. This study focused on the comparison of only two different but relatively large screen sizes. Also, in contrast to our study, the 40" screen was a composite screen with two displays.

In the study of Bi and Balakrishnan [5] on the other hand, a large display wall (16' x 6') was preferred for daily work tasks without touch. For the one week study their participants switched from their usual single or dual desktop display to a large high-resolution display wall. Participants claimed the wall to help them e.g. with multi-window tasks and rich information tasks. In this study, users that used single or dual desktop displays before starting the study, switched for one week to a wall display.

For 3D interaction, a larger screen can have a positive impact. In their cognitive comparison of 3D interaction in front of a wall (4m x 3m) vs. a desktop display (15"), Tyndiuk et. al. [14] showed that the wall had a positive effect on performance for manipulation and naive travel tasks. Although not all users benefited to the same degree from the wall.

The first study to compare performance and ergonomics of touch surfaces was done by Bachynskyi et. al. [3]. In their Fitts' Law experiment they used Motion Capture to compare pointing tasks on five different sized touch surfaces. The displays were set up in different orientations and with specific postures of the participants. They showed that tablet and laptop were poor in performance, while the larger displays (public display, tabletop) as well as the smallest display (smartphone) were high or medium in performance. Though they compared different screen sizes and their impact on performance, postures varied for the displays.

Overall, prior research has not examined the role of display size for touch interaction in a focused way. While there is some evidence that larger size negatively impacts performance [16, 2], other studies suggest that a large size is seen positive for specific applications [5, 14]. The most closely related study suggest that only the tablet/laptop size is poorer in performance than other sizes (smaller or larger) [3]. Our study focuses on comparing screen sizes only, keeping all other factors equal (e.g. user standing for all screen sizes).

3   User Study

For this study, each task required the participants to do translation movements using touch (see Figure 1). Participants were instructed to perform a straight dragging motion between two circles. The start area was shown with a green circle in the center labelled "START". An orange circle marked the target area. As soon as the start circle was touched, it turned grey and the outline of the target circle turned green. The position of the finger was shown with a cursor. The user was instructed to drag the cursor inside the target area. A trial was successful if the cursor was inside the target circle when lifting the finger. Success or failure were signalled by sound and visually with a coloured outline (green/red) around the target circle. Note that it is necessary to differentiate between cursor and actual finger tip position because of latency, all touch systems have a significant amount of lag [10].
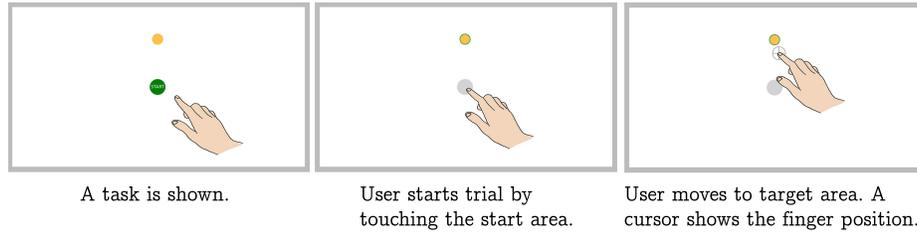

| A task is shown. | User starts trial by touching the start area. | User moves to target area. A cursor shows the finger position. |

Fig. 1: Translation movement task

We tested three display sizes in our study, resulting in the following three conditions (Figure 2):

a. condition SMALL: 13.5" tablet (3000×2000 pixels)
b. condition MEDIUM: 28" monitor (4500×3000 pixels)
c. condition LARGE: 69.5" monitor (1920×1080 pixels)

Fig. 2: Setup for conditions (a) SMALL, (b) MEDIUM, and (c) LARGE.

For an analysis in multiple directions and with different IDs, target circles were computed with three possible distances/amplitudes (17%, 23%, 30% of the

display height), four directions (0°, 90°, 180°, 270°) and three different widths (see Table 1). The distances were calculated using the display height. The widths were computed based on the displays screen resolution and pixel density to create identical ID values. Each configuration was repeated three times.

Target Width			Ampl.	ID	Target Width			Ampl.	ID	Target Width			Ampl.	ID
S	M	L			S	M	L			S	M	L		
12px	18px	9px	17%	4.5	30px	45px	21px	17%	3.3	60px	90px	42px	17%	2.4
12px	18px	9px	23%	4.9	30px	45px	21px	23%	3.6	60px	90px	42px	23%	2.7
12px	18px	9px	30%	5.3	30px	45px	21px	30%	4.0	60px	90px	42px	30%	3.1

Table 1: All combinations of target width (in px) and amplitude (in percentage of display height) with respective ID values for the three conditions S: SMALL display, M: MEDIUM display and L: LARGE display

Participants: We had 16 participants in the study (5 females). The average age was 25.6 years (SD = 7.8). The average body height was 175 cm (SD = 9.9), the average arm length 71 cm (SD = 4.1). All participants were right-handed and had experience using touch displays. Most were students from Computer Science or Design.

Apparatus: The study was conducted under lab conditions. Participants received written introductions. For condition SMALL a Microsoft Surface Book (13.5") was used, a Microsoft Surface Studio (28") for condition MEDIUM and an Iiyama ProLite 70 (69.5") for condition LARGE. The center of each display was positioned at the same height of 140 cm. To record the participants' interactions two cameras were set up on tripods at the sides of the main area.

Study Procedure: Each participant filled out a questionnaire for background information and was given a short briefing for the study. For each condition, the following steps were repeated: The participant performed a short training for the particular screen where each configuration was tried once per direction. After training the actual trials were conducted and measurements were taken. Then, a questionnaire was given to the participant, asking how the task was perceived in this condition regarding factors like difficulty or precision. Each participant was given the opportunity for open comments. Finally, the participant had a three-minute break to avoid fatigue.

After finishing all three conditions, the participant was given a post-test questionnaire comparing the different screen sizes. Also, the participant had the chance to revise previous answers in the questionnaires on each condition. The order of conditions was counterbalanced, i.e., it changed after each participant. For each condition, 36 configurations with three repetitions were executed. For the three conditions this resulted in $3 \times 36 \times 3 = 324$ trials per person. Therefore,

with 16 participants a total of $324 \times 16 = 5{,}184$ trials were recorded. Each session took about 30-45 min to complete.

Measurements: Several aspects were measured in this study. In a first questionnaire data such as age, gender, body height, and arm height as well as technical background and profession was collected. Further questionnaires regarding the display sizes were used to collect subjective data to measure joy, effort, fatigue, difficulty, perceived speed, and perceived precision. Based on the objective time measurements, effective throughput, and error rate were computed.

4   Results

All results were evaluated using a one-way between-subjects ANOVA and post-hoc t-tests using the Bonferroni method.

Speed: To make a fair comparison of speed across different display sizes we calculated speed in in terms of (% of display)/ms. Also, to compare different distances we normalized the paths from start to target to a unit interval of [0,1] and looked at mean speed times along this normalized path (Figure 3). We found significant results at the p<.005 level comparing the three display sizes (F(2,45)=7.51). The post hoc comparision showed that the maximum average speed was higher for the SMALL display than for the LARGE one. The SMALL display had a maximum average speed of 0.25 (SD = 0.051), while the maximum average speed for LARGE was 0.19 (SD = 0.048). No significant results were found comparing the MEDIUM display (M = 0.21, SD = 0.034) with the other displays.

Fig. 3: Average speeds in %/ms at relative distances to the target.

Error Rate: Errors are trials where the final touch point was outside the target area. We found significant results at the p<.01 level (F(2,45)=6.956). The post-hoc test indicated that the error rates of movements on the tablet (M = 15.62%, SD = 6.96) were significantly higher than on the MEDIUM monitor (M = 7.18%, SD = 8.09) or on the LARGE display (M = 8.09%, SD = 5.85) as shown in Table 2. There was no significance between MEDIUM and LARGE.

Performance: Performance was measured as Effective Throughput (TP) [13]. The Index of Difficulty (ID) was calculated using the Shannon Formulation. In

	SMALL	MEDIUM	LARGE
Error Rate	M=15.62%, SD=6.96	M=7.18%, SD=8.09	M=8.09%, SD=5.85
Perceived Difficulty	M=2.00, SD=1.32	M=1.13, SD=0.96	M=0.88, SD=0.72
Perceived Precision	M=2.06, SD=0.85	M=3.06, SD=0.93	M=3.00, SD=0.82

Table 2: Error Rate, Perceived Difficulty and Perceived Precision. Mean and SD Values.

our study the SMALL display had TP of 4.87 (SD = 1.14). TP of the MEDIUM monitor was 5.15 (SD = 0.96), while it was 4.63 (SD = 1.33) for the LARGE display. None of the differences between the performances of the displays were significant (F(2,45)=1.55, p=0.223).

Questionnaire: In our questionnaires we asked about perceived joy, effort, fatigue, difficulty, speed, and precision. Asking the same questions for each display, the subjects gave ratings on a 5-point Likert scale. The answers were mapped to values from 0 to 4. For analysis we treated these values as interval values which is common practice. No significant results could be found for joy, effort, fatigue, and speed. Regarding difficulty (see Table 2), significant results could be found at a p<.01 level (F(2,45)=5.289). The SMALL display (M = 2.0, D = 1.32) was perceived more difficult to use than the LARGE screen (M = 0.88, SD = 0.72). Also significant differences at a p<.01 were found regarding the perceived precision (F(2,45)=6.67). The SMALL screen was perceived less precise (M = 2.06, SD = 0.85) than the MEDIUM (M = 3.06, SD = 0.93) or the LARGE screen (M = 3.0, SD = 0.82) as can be seen in Table 2.

5   Discussion

Which screen size is best for which purpose? Our results give a differentiated picture.

Small Screen (tablet, notebook): Even though a small display (13.5" tablet) has a higher average speed than a very large (69.9") display, our results show that it is problematic objectively and subjectively. Objectively, it has the highest error rate of the three devices. Occlusion may be one reason for the high error rate. As the users move their hand to the target, both their hand and their fingertip cover the target area. As described by Vogel and Balakrishnan [15] occlusion is likely to increase errors. Subjectively, the small display is perceived both difficult to use and imprecise, which corresponds with the high error rate. This result confirms findings by Bachynskyi et. al. [3] that tablets and small notebooks yield a poor performance for touch input.

Medium Screen (desktop monitor): The 27" desktop monitor performed quite well. Regarding error rate, subjective precision, and difficulty it outperformed the smaller display (13.5") while is is not significantly slower than the small display.

Large Screen (smartboard, touch wall): While a large display was claimed to be too difficult to use in the study of Anslow and Wong [2], it performed quite well for translation movements in our study. Not the large display, but the small display was found to be difficult to use. Our findings are in accordance with studies that positively assess large screens in terms of performance [3] and for concrete applications [5, 14]. However, it is significantly slower compared with the small display (13.5"). Also, it is well known that touch interactions on large displays cause fatigue [1].

Comparison: Overall a small display performs worst regarding errors, difficulty, and precision, but may be the best form factor for many applications, e.g. regarding mobility. Large displays on the other side are better in terms of display space and for the use in public scenario setups. Bachynskyi et. al. [3] have shown that though the performance was poor for the tablet or the laptop, it was suitable for long-term use, while other displays like the larger public display have not been suitable for long-term use, even though their performance was high. The input method seems to influence how well a display size performs. While Wang et. al. [16] found no significant difference in accuracy when using a mouse input device, there is an impact when using touch. A small display size (tablet) leads to a higher error rate and thus is less accurate. Regarding the subjective difficulty, our results differ from the research of Anslow and Wong [2]. While they showed that a larger display was perceived more difficult to use, we show that smaller displays are perceived most difficult to use, with no significant difference between medium and large displays. This contrast is even more surprising as our size difference is much larger (27" against 69.5") than that of Anslow/Wong (27" against 40"). So we assume that the difference hinges on their specific scenario but we conclude that a larger display has no general disadvantages in terms of performance but certainly in terms of fatigue [1].

6   Conclusion and Future Work

We explored how display size impacts performance of touch translation movements. Three display sizes were tested: small (13.5" tablet), medium (28" monitor), and large (69.5" screen). We analyzed performance, error rate, speed, and subjective values such as effort, joy, difficulty, fatigue, and precision. The smallest sized display performed worst. Though its speed was higher than for the largest display, its error rate was highest and it is perceived both imprecise and difficult-to-use. Small displays are therefore suited for scenarios that require high mobility and speed (like games). The medium and the large screens, in contrast, performed quite well. Since very large screen cause more fatigue the medium screen may be the "golden middle" for the majority of applications.

For the future, we plan to use motion capture analysis to further investigate the correlation between performance and human movement [3]. Ultimately, we want to find a model of human movement that explains various findings on display size, angle and mobile device grips and its relation to measurable performance.

References

1. Al-Megren, S., et. al. (2015): Comparing Fatigue When Using Large Horizontal and Vertical Multi-touch Interaction Displays. In: Proceedings of Interact, pp. 156-164 https://doi.org/10.1007/978-3-319-22723-8_13
2. Anslow, C., Wong, W. (2017): Effects of the Display Angle and Physical Size on Large Touch Displays in the Work Place. In: Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces (ISS), pp. 318-323 https://doi.org/10.1145/3132272.3135080
3. Bachynskyi, M., et. al. (2015): Performance and Ergonomics of Touch Surfaces. A Comparative Study Using Biomechanical Simulation. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 1817-1826 https://doi.org/10.1145/2702123.2702607
4. Baudisch, P., et. al. (2003): Drag-and-Pop and Drag-and-Pick. Techniques for accessing remote screen content on touch- and pen-operated systems. In: Proceedings of Interact, pp. 57-64
5. Bi, X., Balakrishnan, R. (2009): Comparing Usage of a Large High-Resolution Display to Single or Dual Desktop Displays in Daily Work. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI), pp. 1005-1014 https://doi.org/10.1145/1518701.1518855
6. Fitts, P. M. (1954): The information capacity of the human motor system in controlling the amplitude of movement. In: Journal of Experimental Psychology, 47.6, pp. 381-391 https://doi.org/10.1037//0096-3445.121.3.262
7. Hoggan, E., et. al.(2013): Multi-Touch Pinch Gestures: Performance and Ergonomics. In: Proceedings of the 2013 ACM International Conference on Interactive Tabletops and Surfaces, pp. 219-222 https://doi.org/10.1145/2512349.2512817
8. Hoggan, E., et. al. (2013): Multi-Touch Rotation Gestures: Performance and Ergonomics. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 3047-3050 https://doi.org/10.1145/2470654.2481423
9. Lehmann, F., Kipp, M. (2018): How to Hold your Phone When Tapping. A comparative study of performance, precision, and errors. In: Proceedings of the ACM International Conference on Interactive Surfaces and Spaces (ISS), pp. 115-127 https://doi.org/10.1145/3279778.3279791
10. Ng, A., et. al. (2012): Designing for low-latency direct-touch input. In: Proceedings of the 25th annual ACM symposium on User interface software and technology, pp. 453-464 https://doi.org/10.1145/2380116.2380174
11. Nguyen, Q., Kipp, M. (2014): Orientation Matters: Efficiency of translation-rotation multitouch tasks. In: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 2013-2016 https://doi.org/10.1145/2556288.2557399
12. Nguyen, Q., Kipp, M. (2015): Where to Start? Exploring the efficiency of translation movements on multitouch devices. In: Proceedings of the Interact, pp. 173-191 https://doi.org/10.1007/978-3-319-22723-8_15
13. Soukoreff, R., MacKenzie, I. (2004): Towards a standard for pointing device evaluation, perspective on 27 years of Fitts' Law research in HCI. In: International Journal of Human-Computer Studies, 61.6, pp. 751-789 https://doi.org/10.1016/j.ijhcs.2004.09.001
14. Tyndiuk, F., et. al. (2005): Cognitive Comparison of 3D Interaction in Front of Large vs. Small Displays. In: Proceedings of the ACM symposium on Virtual reality software and technology (VRST), pp. 117-123 https://doi.org/10.1145/1101616.1101641