Exploring Fitts' Law for Touch Screens of Various Sizes

B.S. Greve[1], P. Kostic[2] and H.I. van Maarle[3]

[1]Department of Computer Science, Utrecht University, Utrecht, The Netherlands. [1]b.s.greve@students.uu.nl

[2]Department of Computer Science, Utrecht University, Utrecht, The Netherlands. [2]p.kostic@students.uu.nl

[3]Department of Information Science, Utrecht University, Utrecht, The Netherlands. [3]hasse.whitedove@gmail.com

Abstract

Fitts' law is an accurate measure of how difficult it is to make precise hand movements from point A to point B. This research examines whether Fitts' law holds for touch screens of various sizes (an iPhone 6+ and an iPad air). For the experiment two touch input devices, pen and finger, were used. 20 students of the university of Utrecht participated in the test and were asked to drag a black box towards a randomly placed green box 20 times in a row. No significant proof that Fitts' Law is applicable to touch screen has been found in the collected data.

Keywords: *Fitts' Law, Touchscreen, Stylus Pen Input, Finger input, Screen size*

Introduction

The Human processor model (Card, Moran, & Newell, 1986) or MHP (Model Human Processor) is a cognitive modeling method used to calculate how long it takes to perform a certain task. In this model, the time it takes to perform a micro-movement is equal to the "feedback-loop" / "reaction time". This "feedback-loop" / "reaction time" includes observations with perceptual (visual) processors, evaluations of positions (cognitive processing) and then performing a new movement as a response to the evaluation. According to the MHP model, the average perceptual processor cycle time is 100ms, the cognitive processor cycle time is 70ms and the motor processor cycle time is 100ms. Adding these together results in a 240ms time for a micro movement.

In 1954 Paul Fitts introduced what became known as Fitts' Law, a model to predict the time required to rapidly move to a target area. Fitts conducted multiple one dimensional tests and proposed a formula. This formula stated that the time required to rapidly move to a target area is a function of the "distance to" and the "size" of the target. Thus, Fitts' model mathematically describes a movement consisting of several controlled micro-movements. In HCI, Fitts' Law is typically defined as:

$$T = RT + MT = a + b\ log_2(\frac{2A}{W} + c)$$

where T is the time to complete the task, RT is reaction time, MT is the average time taken to acquire the target, a and b are empirical constants determined through linear regression, A is the distance from the starting point to the center of the target, W is the width of the target measured along the axis of motion and c is a constant which is either 0, 0.5, or 1, depending on the specific environment. Note that the logarithmic function implies that a small increase in size or distance is much more effective for small targets than large targets.

Originally, Fitts tested his law with three versions, none of which were digital (Fitts, 1954). Since then however, it has been shown that Fitts' law, with some minor adaptations to improve its accuracy in a 2D environment, also works when applying it to the movement of a mouse on a computer screen (MacKenzie and Buxton, 1992). Although the more traditional mouse as an input device still exists, a lot of new input devices have been introduced into the daily life of society. Touch screens on smartphones and tablets are very common and for touch screens there are various input methods such as a finger or various touch pens. Screen sizes have also changed a lot since then, with bigger wide aspect-ratio screens and small smartphone screens with high pixel density. Bi, Li and Zhai have suggested that additional adaptations to Fitts' law are necessary when applying the law to touch screens (Bi, Li and Zhai, 2013). However, their version was used for 1D tests rather than 2D. Additionally, different screen sizes are not accounted for.

This paper focusses on testing Fitts' law on touch screens of varying sizes (an iPad Air and an iPhone 6+) and with varying touch input devices: a finger and a pen. In this paper the focus will more specifically lie on testing Fitts' Law amongst students of Utrecht University. The question that will be answered is: does Fitts' law still hold for touch screens of varying sizes with varying input devices when tested on students? This choice was made due to resource constraints as well as the fact that these are the people that are most likely to spend their working life involved with touch screens and therefore it is most important to see if the law upholds for this group.

As a series of tests often shows the participants learning more about the test and progressively getting better at it (Volker Benndorfy, Holger A. Rauz, and Christian Solchx, 2014), several measures of combating learnability are used in this experiment. These involve a certain amount of randomization and systematically changing the order in which the different tests are taken so they all have the same amount of learnability bias.

The rest of this paper has been divided into three large sections, making the whole consist of four. The first section of the paper will describe the hypotheses and how they are attempted to be proven by elucidating the design of the tests and what data they will be used to gather. This will be followed by an explanation of the statistical analysis used on this data to come to both the conclusions made at the end of the paper and the certainty of these conclusions. Lastly, potential reasons for the certainty rates are discussed, as well as potential further research relating to this paper.

Methods

Previous methods of testing Fitts' law regarding computer interfaces have involved using a cursor to select a given area with a certain width as rapidly as possible (MacKenzie, 1995), an adaptation into two dimensions featuring a goal the shape of a box with a given width and height, and a drag-and-drop approach, with two boxes in a 2D field of which one has to be dragged over to the other. One commonality of these methods was the classic combination of a computer mouse and a CRT monitor. To adopt these approaches to testing Fitts' law to the modern user interface experience, the decision was made to reuse the drag-and-drop experiment (MacKenzie, 1995) with the nowadays ubiquitous touch screen input of smartphones and tablets. Specifically, both a smartphone and tablet will be tested with both touch and stylus pen as input. As such, MacKenzie's experiment and analysis was adopted and slightly adapted to our needs.

Hypotheses

H1: *Touch input can be accurately modeled using Fitts' Law.*

H2: *There is correlation between MT and ID for at least one input device.*

H3: *Movement time is dependent of device or input*

H4: *Error rate is dependent of device or input*

The hypotheses are ranked by importance from 1 to 5.

Participants

Twenty students from Utrecht University (13 women and 7 men) ranging in age from 18 to 27 years old, were randomly selected and voluntarily participated in this experiment after approaching them on campus and asking for participation after a quick overview. All but one of the participants used touch input on a daily basis. To further increase chances of finding participants with a relatively high tech literacy, the random selection of subjects took place at and around the STEM faculties of a university campus.

Materials

All tests were performed using an *iPhone 6 Plus* and an *iPad Air 1* (both running iOS 10.3.2) with input through touch and a stylus pen. The visual output of the *iPhone 6 Plus* was a 5.5" IPS display with a resolution of 1080 by

1920 pixels (401 ppi). The visual output of the *iPad Air 1* was a 9.7" IPS display with a resolution of 2048 by 1536 pixels (264 ppi). The stylus pen was a rubber tipped ballpoint pen that emulates a user's finger.

We used our own software (see SOM) to record the dragging coordinates, dragging time and task completion of the participants. Records were saved in a CSV format to the device's HTML5 local storage and retrieved afterwards using the Safari Debugger. Both devices were connected to the internet during all tests.

A questionnaire with both multiple choice and open questions was constructed and used to acquire more information regarding the skill levels of the various subjects (see Appendix A). These questions relate to the prior experience of the subject with a variety of input devices so these relative differences can be accounted for in statistical analysis of the data generated by testing with the applet, but also leaved open the possibility of the subject to mention any other factors that may have impacted their performance, such as their current mental state or physical ailments. The questionnaire was conducted on an *iPad Air 1* with the results recorded on a spreadsheet (see Appendix B).

Design

The experiment used a 2 x 2 x 2 x 2 between-subjects design. Controlled variables were device (two levels), success (two levels; TRUE and FASLE), target amplitude (two levels), target width (two levels). Dependent variables were movement time (MT), error rate (calculated from the recorded coordinates), and the index of performance ($IP = 1 / b$, from $MT = a + b\,ID$; ID calculated using the Shannon formulation). Movement time was measured from when the participant's touch location is within the movable box (starting a dragging movement) to when the participant's finger releases from the screen (ending a dragging movement).

The experiment was run by having 42 dragging rounds on each device and input method, resulting in a total of 84 rounds.

Phone touch -> phone pen -> tablet touch -> tablet pen
Phone pen -> tablet touch -> tablet pen -> phone touch
Tablet touch -> tablet pen -> phone touch -> phone pen
Tablet pen -> phone touch -> phone pen -> tablet touch

One test for the *iPhone* specifically includes 5 tasks at 522px distance with 220px goal size, then 5 tasks at 870px distance with 220px goal size, followed by 5 tasks with 522px distance and 112px goal size, and lastly 5 tasks with 870px distance and 112px goal size. The task was similarly structured for the tablet and scaled to its ppi, resulting in similar real-world widths, heights and distances present on the screen

Procedure

Every participant undergoes the 84 rounds divided over a series of four tests, one for each combination between both input methods and mobile devices. To keep the numbers equally valid across the tests, every possible order to take the tests in was carried out the same amount of times. As such, the learnability across multiple tests does not disproportionately affect any of the input methods or screens, allowing at least that small portion of learnability to be rendered irrelevant.

The operation of the devices and the requirements of the rounds were demonstrated and explained to each participant before beginning. Six warm-up rounds were given on the starting device and discarded from data collection by refreshing the Safari web browser.

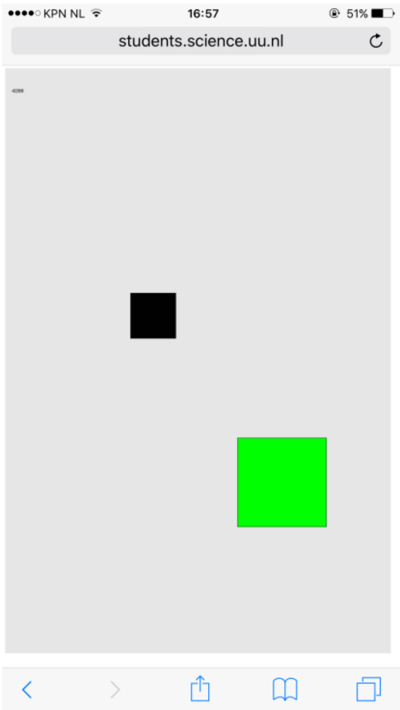

Figure 1 – An example of the initial round on the *iPhone 6 Plus.*

Each round was a “drag-and-drop” task in the Safari web browser of each device. The applet itself consists of a canvas on which two boxes are drawn upon runtime (see Figure 1). One is black and starts in the middle of the canvas, while the other is green. The first of these is the interactive movable box that is to be dragged around by the test subject, and the second is the goal to which the first is to be dragged and released. The smallest instance of the test is called a “round” and is defined as either succeeding or failing to drag the black box onto the green box without letting go, upon which the next round is started. The start of a new round has the green box removed from the screen. If the previous round was successful the black box is left where it was let go of, within the green area, but if the opposite case is true and the round was failed, its position is reset to that of the removed green box so the next round can be attempted unphased by the previous loss.

The questionnaire was administered after the subjects had done all four tests as reminding people of stereotypes before they perform a test might influence how well they perform (Shih, Pittinsky, & Ambady, 1999).

Results

To properly interpret the relationships between the data points gained from the different tests, two-way ANOVA was employed to examine the main influence of both input methods and devices on the movement time of the participants, but also to establish the interaction between the input methods and devices themselves.

Movement Time

The mean MT when taking all tests into account was 452.1613ms, and the mean for the same test with only successful drags was 495.1804ms. Mean distribution by tests can be found in Figures 2 and 3.

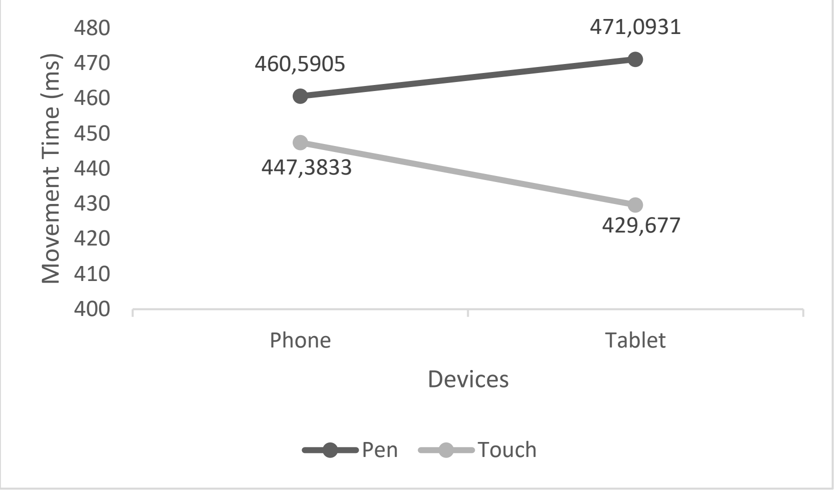

Fig 2. - Mean time per test.

Figure 2 indicates that the mean MT for touch input were significantly faster than pen input ($F_{1,1677} = 5.9725$, $p = 0.0146$). Figure 2 further implies that phones have a lower MT than tablets when it comes to using a pen ($F_{1,1677} = 0.1048$, n.s.), but a higher MT when it comes to using touch ($F = 0.1048$, n.s.).

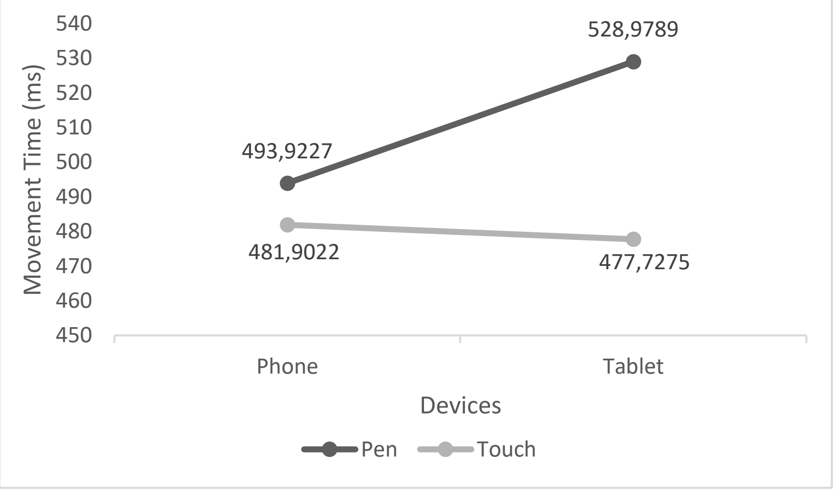

Fig 3. - Mean time per successful test.

When comparing Figure 2 and Figure 3, it is interesting to note that removing failed tasks leads to significantly higher MT across the board. Figure 3 further strengthens the notion that using a stylus pen is significantly slower than using touch input ($F_{1,1404} = 8.1374$, $p = 0.004$). However, we noticed small statistically significant interaction between the effects of Input and Device on the MT ($F_{1,1404} = 3.3680$, $p = 0.06669$). Since the interaction effects almost approaches significance of $p < 0.05$, the Input effect cannot be generalized for both devices when removing failed rounds.

Errors

An error was defined as dragging the black box outside of the green box target. Thus, premature releases while dragging the black box or accidental taps on the black box itself resulted in failed rounds. The mean error rates for all tests were 16.2%. Error rate distribution by test can be found in Figure 4, which only shows tablet having significantly higher error rate for both input methods ($F_{1,1676} = 27.0576$, $p < 0.001$).

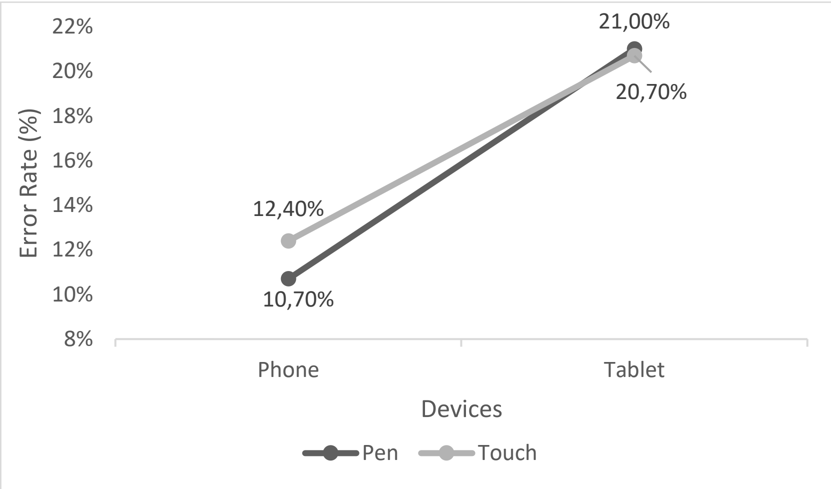

Fig. 4 – Mean error rate per test.

Regression Analysis

A linear regression was calculated using MT as the dependent variable and ID as the independent (criterion) variable, which was computed using the Shannon formulation. The results are shown in Table 1.

Table 1 - Regression Analyses per tests

Device	Input	R^2	*SE* [a] (ms)	Pr(>|t|) [c]	Regression Coefficients		
					Intercept (ms)	Slope (ID) (ms/bit)	*IP* (bits/s) [b]
Phone	Pen	0.1051	20.95	1.06e-06	245.29	105.28	9.5
	Finger	0.1056	24.22	3.94e-07	199.13	126.51	7.9
Tablet	Pen	0.1814	23.15	1.67e-09	136.59	147.00	6.8
	Finger	0.1523	24.50	6.08e-08	133.74	138.56	7.2

[a] standard error of Slope (ID) estimate
[b] *IP* (Index of performance) = 1 / b, where b is the slope
[c] p-value for the independent variable (ID)

There were low correlations across the board between movement time (MT) and the index of task difficulty (ID) for all device-input combinations ($R^2 < 0.2000$). Performance indices (IP), ranging from 6.8 bits/s to 9.5 bits/s, were bigger than those found by MacKenzie (1995) and approaching those of Fitts (1945). The phone outperformed the tablet using both input methods, with slight differences for touch input.

Because we were interested in the low correlations, all regressions for all four tests were plotted (Figures 5, 6, 7 and 8).

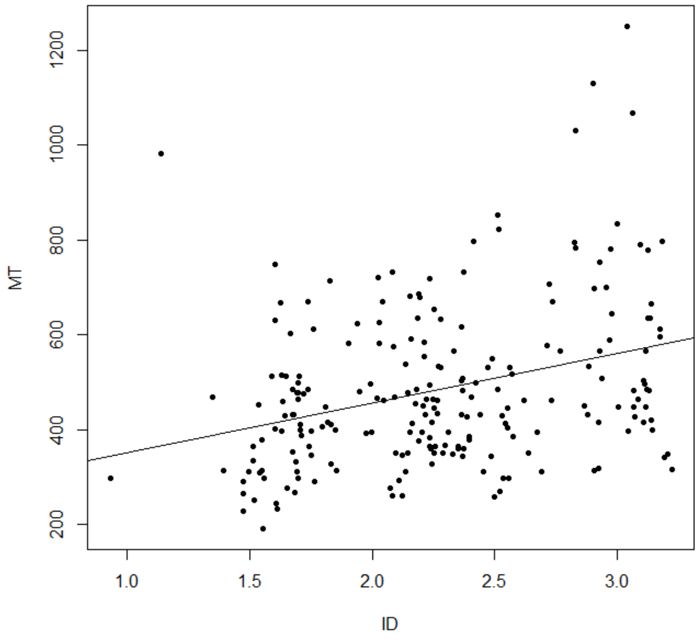

Fig. 5 – Scatter plot and regression line for Phone – Touch. MT = 199.13 + 126.51 ID

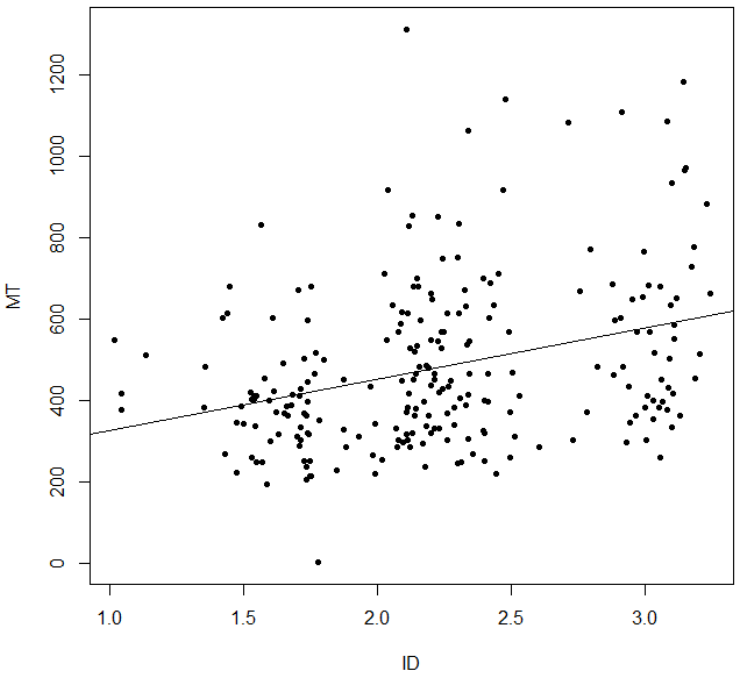

Fig. 5 – Scatter plot and regression line for Phone – Pen. MT = 245.29 + 105.28 ID

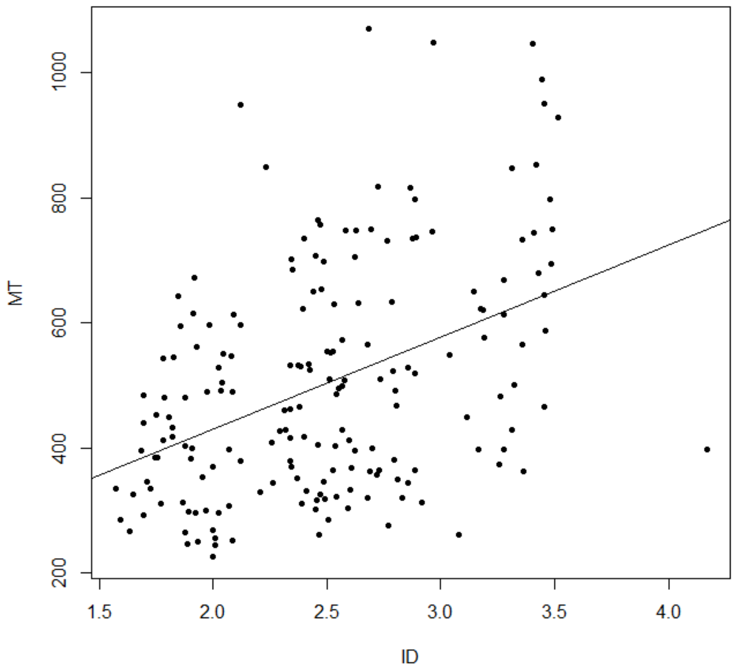

Figure 7 – Scatter plot and regression line for Tablet – Pen.MT = 136.59 + 147 ID

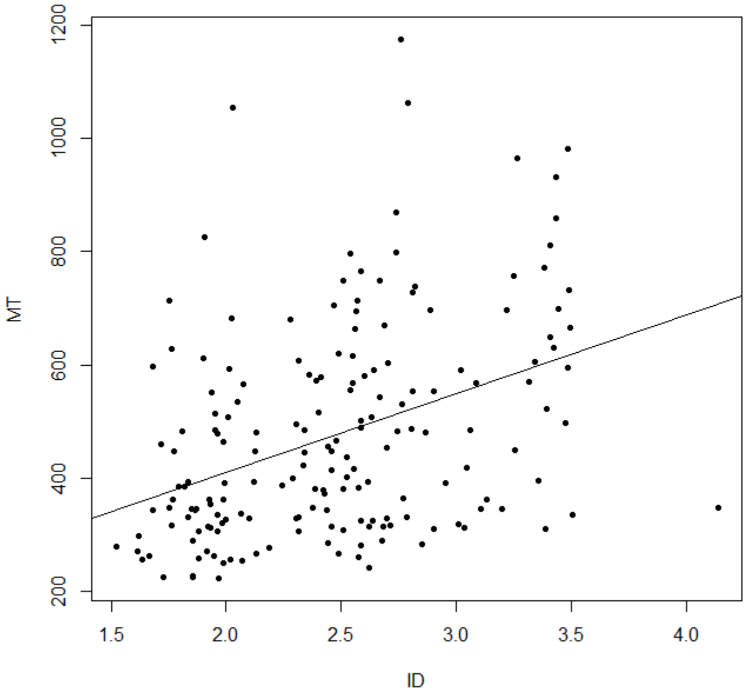

Fig. 8 – Scatter plot and regression line for Tablet – Pen. MT = 133.74 + 138.56 ID

As can be seen in figures 5, 6, 7 and 8 and the p-values in the previous section, the data points recorded by the tests are sufficiently scattered to make a linear regression line look rather unconvincing. This model's predictive power is lacking when compared to MacKenzie, I. S. (1995) fig. 4, besides suffering from the problem MacKenzie mentions of the intersection with the y-axis not being at the origin, implying an unintuitive relationship between ID and movement time.

Discussion

Summary

This research paper has attempted to replicate the prediction power of Fitts' law in the context of modern touch screen devices. Four tests were carried out consisting of every combination between the devices of smartphone and tablet, and the input methods of touch and stylus pen.

For the experiment 20 students of the university of Utrecht were asked to perform 4 tests on a custom-made application. Two per test device, once using their finger as an input method, once using a special touch pen. During each test, they were asked to move a small black square into a green square that varied in size, distance from the black square and location. The order in which the students took the test was varied to have an equal distribution of learned skills over the various tests. Per test the students got 21 green boxes that appeared after the previous green box was either successfully hit or missed. Afterwards students were asked to fill in a short questionnaire that asked them about their experience with various input devices, their age, their gender and whether there was anything else (alcohol, etc.) that might impact their speed. This was intentionally done afterwards as reminding people of stereotypes before they perform a test might influence how well they perform (Shih, Pittinsky and Ambady, 1999). Then the test results were compared to Fitts' Law using R to determine if Fitts' Law held true for these results

H1: *Touch input can be modeled using Fitts' Law.*

Mostly rejected. Regression analysis showed that we were unable to accurately model Fitts' Law using our research method. However, different methods might yield better results. These have been included in the future work section of the discussion.

H2: *There is correlation between MT and ID for at least one input device.*

Rejected. The summary of the regression analysis showed low R^2 values, indicating low correlation for all input devices.

H3: *Movement time is dependent of device or input.*

Partially rejected. Our results showed that, for touch input, movement time was significantly faster for pen input. However, our results showed no statistical significance for device.

H4: *Error rate is dependent of device or input.*

Not rejected. Our results showed that a tablet has significantly higher error rate for both input methods. However, our results showed no statistical significance for input.

Method Rationale

To achieve our initial research goal, a simple applet was written in Processing.js. Processing.js specifically was chosen for its quick prototyping capabilities, multi-platform compatibility and browser-based nature. All these factors made it simple to write and iterate on the applet, and get it running on all available test devices.

To calculate if Fitts' law still holds for the modern devices the test was carried out on, Shannon's original formula (Shannon, C. E., & Weaver, W., 1949) was used:

$$ID = log2(2A / W + 1)$$

With ID = Index of Difficulty in bits, A = Amplitude in inches and W = width of target in inches. Shannon's formula specifically was used because Fitts' formulation of the law tends to fail for sufficiently small ID values (MacKenzie, I. S., 1989). As the first set of rounds has a large goal size with a low distance, it was decided to use Shannon's instead.

To say that Fitts' law still holds one would have to show that the relationship between the ID of a task and the time taken to accomplish the task is a linear one. Considering the nature of the tests, A is easily available as it was one of the key changing variables in the series of tasks. A simple R script is used to calculate W using the assumption of a straight drag from the black box to the green one. This assumption is made so that the computation of W in a 2D field can be reduced to a 1D problem, greatly reducing the problem in complexity. The script calculates the intersection points between the line formed by using the middle of the two boxes as anchor points and the green goal box, and then calculates the distance between those two intersection points (see Figure 9). The code of this script can be found in appendix C.

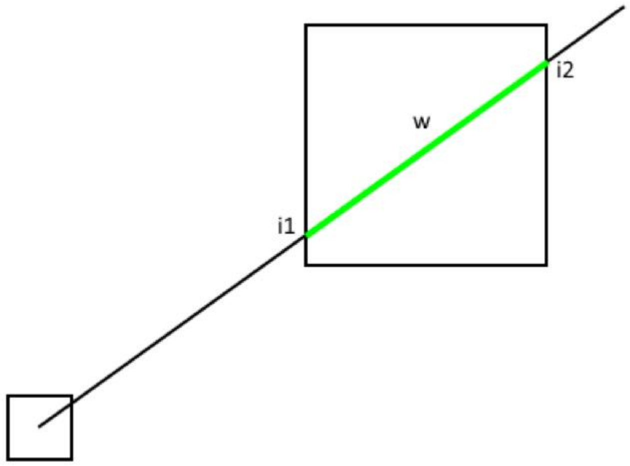

Fig. 9 – Illustration of the acquisition of W.

To reduce the learnability of both the whole test and the individual rounds, an element of randomness is introduced in the placement of the green boxes. Because the distance between the start and end of the interaction is an important facet of Fitts' law, it was necessary to ensure control over this distance so that multiple time measurements over the same distance could be recorded. Due to this fact, the only possible positions of the next green box are those on the circle of radius d where d is the distance selected for a given sequence of rounds. As such, every round a random angle is chosen at which the next green box should appear at d distance from the black box. If this position turns out to be outside the canvas, a new random angle is attempted. This system ensures that there is a possible position for the next green box as long as the black box is within d distance of the canvas. Considering it should never be able to leave the canvas, this leaves the applet with a lot of random positions to choose from for any d small enough to not fully envelop the canvas in its circle.

The applet is structured in such a way that rounds can be added simply by being provided a d at which they are to place their respective green box relative to the previous round, and the size of the green goal box. In this way, a full test can be constructed by coming up with an order in which to test various distances.

To keep the scale of this study manageable and specific enough to be useful, it was elected to apply the certain constraints to the sample of testers. Because tech literacy is on the rise, it can be assumed that most future users of interfaces will have grown up with a variety of modern computer UIs. To keep this study relevant for the future it was decided to create the focus group within such an age that they are likely to have grown up with easy access to various modern digital devices.

Limitations

The research done in this paper has various limitations. The first is the small sample size. Utrecht University has a population of 30374 students. For a confidence level of 95% and a confidence interval of 5, 379 students would be necessary. For this calculation, the method of Krejcie and Morgan was used (Krejcie and Morgan, 1970). Unfortunately, due to time and resource constraints it was not possible to choose a bigger sample size.

Another limitation is the narrow scope of the research, it focuses only on young people that study in the Netherlands, this limits the generalization of the research results. Due to resource constraints, the research was not done in a lab, this meant that the environment was often noisy and students were sitting or standing in various positions with alternatingly the test devices in their hands or on a table. Effort was made however to ensure that there were no problems with sunlight reflecting off the screens. The nature of the test devices was such that if the test subjects started dragging at or outside the edge of the screen the canvas would drag along, causing a disruption and thus potential delay.

Another thing that has influenced the test results is the fact that different test subject did the tests at different paces and accuracies. Assumed is that this is because some subjects focused on being fast while others focused on being accurate. Another limitation is the fact that amongst the test subjects there was a high familiarity with touch screens but a low familiarity with touch pens.

Lastly, a significant percentage of data points was lost in calculating the length of the drag-line segment in the goal box due to a rounding error in R. In a series of edge-cases such as the slope of the line being 0 or infinite problems arose that would not be difficult to combat in further research. Thus, a lot of rows got a W' result of NA, which rendered them unusable for further analysis. Thanks to the randomness in the location of the goal these NA values were distributed quite well among the different tests, and should not introduce a major bias in the results. The code of the script that caused this problem can be found in Appendix C.

Future Work

If we look at directly dragging a box on a touch screen from a discontinuous perspective and taking 3D movement into consideration, multiple challenges come into play. For instance, the origin of movement is often the default position of the finger. This varies a lot depending on the device and where the screen is mounted. On a hand-held device, a test subject might hold the device in one hand and use their other hand's index finger to complete the task. Another possibility is either laying the device flat on a table and use their index finger or holding the device in one hand and using that hand's thumb. Another problem arises from the use of a stylus pen, especially the angle of the target box with regards to the relative positioning of the target with regards to the focus of the user. There could be an argument made that the angle of the target and the relative positioning could be the constants a and b used in Fitts Law

To combat these problems as much as possible, the decision was made to keep the location of the black box stable between successful rounds, it stays wherever the participant put it during the last round. This minimizes the necessity of 3D movement between rounds, and ensures the test isn't missing an entire dimension of significant information. However, this remains a naïve approach to measuring touch movement and can be improved by measuring the finger location and subsequent movements in a 3D space.

The data for this study was not adjusted in any way. Any statistical outliers were kept in the data set. In the future, a better model might be found by discarding certain outliers, leading to a more representative and statistically significant regression line. The removal of these outliers could be motivated by using data from the questionnaire or recording the situation in which the subject takes the test to search for irregularities.

Another possible adjustment that could be made to the data is using the questionnaire to add weights to certain results based on the participant's level of experience with the given input method and device. More research into just how these weights could be implemented would have to be conducted.

Conclusion

In this paper an experiment has been done to test if Fitts' Law would still hold for touch screen devices of varying screen sizes and varying input methods. The hypothesis that touch input can be modeled by Fitts' Law was mostly rejected. The hypothesis that there is a correlation between MT and ID for at least one input device was rejected. The hypothesis that MT is dependent on device or input type was partially rejected. The hypothesis that the error rate is dependent of device or input has not been rejected. The fact that so many hypotheses were at the very least partially rejected should most likely be attributed to the limitations of this research as the research was not performed in a controlled environment and on only very few test subjects. Future research should focus on avoiding these limitations, whether outliers should be taken into acocunt and on modeling movement of the touch input device (finger or pen) in 3D space.

References

Bi, X., Li, Y., & Zhai, S. (2013). FFitts law. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13, pp.1363-1372.

Card, S. K., Moran, T. P., & Newell, A. The Model Human Processor: An Engineering Model of Human Performance. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), Handbook of Perception and Human Performance. Vol. 2: Cognitive Processes and Performance, 1986, pages 1–35.

Fitts, P. (1954). The information capacity of the human motor system in controlling the amplitude of movement. Journal of Experimental Psychology, 47(6), pp.381-391.

Krejcie, R.. & Morgan, D. (1970). Determining Sample Size for Research Activities. Educational and Psychological Measurement, 30(3), pp.607-610.

MacKenzie, I. S. (1989). A note on the information-theoretic basis for Fitts' law. Journal of Motor Behavior, 21, 323-330.

MacKenzie, I. S. (1995). Movement time prediction in human-computer interfaces. In R.M. Baecker, J. Grudin, W.A.S. Buxton, and S. Greenberg (Eds.), Readings in Human-Computer Interaction: Toward the Year 2000 (2nd ed.), p. 483-493. San Francisco, CA, USA: Morgan Kaufmann Publishers, Inc. ISBN: 1-55860-246-1

MacKenzie, I. and Buxton, W. (1992). Extending Fitts' law to two-dimensional tasks. Proceedings of the SIGCHI conference on Human factors in computing systems - CHI '92, pp.219-226.

Shannon, C. E., & Weaver, W. (1949). The mathematical theory of communication. Urbana, IL: University of Illinois Press.

Shih, M., Pittinsky, T. and Ambady, N. (1999). Stereotype Susceptibility: Identity Salience and Shifts in Quantitative Performance. Psychological Science, 10(1), pp.80-83.

Volker Benndorfy, Holger A. Rauz, & Christian Solchx, (2014), Minimizing Learning Behavior in Experiments with Repeated Real-Effort Tasks, Heinrich Heine University Dusseldorf - Dusseldorf Institute for Competition Economics (DICE)

Appendix A – Questionnaire

The questionnaire consisted of 16 questions, as listed below. Beneath the multiple-choice questions are the answers that more than 0 people chose and how many people chose them. The questions were either open questions or they had the following multiple choice options:

Q: Do you use a computer mouse?
No: 5
Yes, daily: 5
Yes, weekly: 6
Yes, monthly: 1
Yes, rarely: 3

Q: For how many years has the above been true? (three, less than one, etc.) Open question

Q: Do you use a trackpad?
No: 3
Yes, daily: 14
Yes, monthly: 2
Yes, rarely: 1

Q: For how many years has the above been true? (three, less than one, etc.)
Open question

Q: Do you use a draw pad?
No: 18
Yes, monthly: 1
Yes, rarely: 1

Q: For how many years has the above been true? (three, less than one, etc.) Open question

Q: Do you use a touchscreen?
Yes, daily: 19
Yes, weekly: 1

Q: For how many years has the above been true? (three, less than one, etc.) Open question

Q: Do you use a touch pen?
No: 19
Yes, rarely: 1

Q: Do you use any other input devices? Open question

Q: For how many years has the above been true? (three, less than one, etc.) Open question

Q: What are you? Open question (youngest participant was 18, oldest 27)

Q: Can you think of anything that might influence your accuracy? Open question

Q: What gender are you?
Female: 13
Male: 7

In what order did you perform the tests?
Open question, to keep track of the order the participants performed their tests in.

Appendix B – Questionnaire results

Timestamp	Do you use a computer mouse? (if yes, for how many years)	Do you use a trackpad? (if yes, for how many years)	Do you use a draw pad? (if yes, for how many years)	Do you use a touchscreen? (if yes, for how many years)	Do you use any other input devices? (if yes, for how many years)	Do you use a touch pen?	What age are you?	Can you think of anything that might influence your accuracy?	What gender are you?
6/1/2017 16:20:28	Yes, monthly, 15	Yes, daily, 7	No, X	Yes, daily, 6	No, X	No	23	No	Female
6/2/2017 12:21:53	Yes, daily, 6	No, X	No, X	Yes, daily, 8	No, X	No	19	No	Male
6/2/2017 12:28:52	Yes, daily, 10	Yes, rarely, 6	No, X	Yes, daily, 7	No, X	No	21	No	Male
6/2/2017 12:32:57	No, 2	Yes, daily, 5	No, X	Yes, weekly, 3	No, X	No	20	No	Female
6/2/2017 12:38:27	Yes, weekly, 15	Yes, daily, 5	Yes, rarely, 4	Yes, daily, 5	No, X	No	22	No	Male
6/2/2017 12:42:53	Yes, daily, 4	Yes, monthly,6	No, X	Yes, daily, 9	No, X	No	23	No	Female
6/2/2017 13:34:20	Yes, weekly, 10	No, X	No, X	Yes, daily, 7	No, X	No	19	No	Female
6/2/2017 13:49:02	Yes, rarely, 18	Yes, 5	No, X	Yes, daily, 7	No, X	No	23	No	Female
6/2/2017 13:54:56	Yes, daily, 18	Yes, daily, 12	No, X	Yes, daily, 10	No, X	Yes, rarely, 10	22	No	Male
6/2/2017 14:01:33	No, X	Yes, daily, 7	No, X	Yes, daily, 8	No, X	No	27	No	Female
6/2/2017 14:05:49	No, X	Yes, daily, 8	No, X	Yes, daily, 8	No, X	No	27	No	Female
6/2/2017 14:10:47	Yes, weekly, 2	Yes, daily, 10	No, X	Yes, daily, 10	No, X	No	24	No	Female
6/2/2017 14:14:17	No, X	Yes, daily, 10	No, X	Yes, daily, 5	No, X	No	24	No	Female
6/2/2017 15:23:56	Yes, daily, 10+	Yes, monthly, 8	Yes, monthly, 5	Yes, daily, 8	Joystick, weekly, <1	No	23	No	Male
6/2/2017 15:29:37	Yes, weekly, 10	No, X	No, X	Yes, daily, 8	No, X	No	19	No	Female
6/2/2017 15:33:26	Yes, weekly, 10	Yes, daily, 1	No, X	Yes, daily, 7	PlayStation controller, weekly, 13	No	18	No	Male
6/2/2017 15:39:04	Yes, rarely, 1	Yes, daily, 8	No, X	Yes, daily, 5	No, X	No	25	No	Male
6/2/2017 15:45:03	Yes, rarely, 5	Yes, daily, 9	No, X	Yes, daily, 8	No, X	No	26	No	Female
6/2/2017 15:50:59	No, X	Yes, daily, 5	No, X	Yes, daily, 4	No, X	No	19	No	Female
6/2/2017 15:54:20	Yes, weekly, 11	Yes, daily, 6	No, X	Yes, daily, 6	No, X	No	20	No	Female

Appendix C – W' function (Language: R)

```
myfunction <- function(startX, startY, endX, endY, boxSize) {
  if(endX - startX == 0){return(NA)}
  slope <- (endY - startY) / (endX - startX)
  if(!is.na(slope)){
        b <- endY - (slope * endX)
        lowerBound <- endY - (boxSize/2)
        upperBound <- endY + (boxSize/2)
        leftBound <- endX - (boxSize/2)
        rightBound <- endX + (boxSize/2)
        rewritten <- 0 - (1/slope)
        if(slope != 0) {
                intersectionV1 <- solve(slope, lowerBound - b, tol = 1e-26)
                intersectionV2 <- solve(slope, upperBound - b, tol = 1e-26)
                intersectionH1 <- solve(rewritten, leftBound - b, tol = 1e-26)
                intersectionH2 <- solve(rewritten, rightBound - b, tol = 1e-26)
                intersections <- rbind(c(intersectionV1, lowerBound), c(intersectionV2,
                                upperBound), c(intersectionH1, leftBound),
                                c(intersectionH2,rightBound))
                df <- as.data.frame(intersections)
                filtersections <- subset(df, V1 > leftBound - 1 & V1 < rightBound + 1 & V2 <
                                upperBound + 1 & V2 > lowerBound - 1)
        w <- sqrt((filtersections[2,1] - filtersections[1,1])^2 + (filtersections[2,2] –
                filtersections[1,2])^2)
        return(w)
        } else {
          return(NA)
        }
   } else {
        return
        }
}
```

Supplementary Online Material

The source code for our recording software can be viewed on Petar’s Github
https://github.com/Snookik/INFOUE

All applets can also be accessed online using the following URLs:
Desktop (development & testing build)
http://www.students.science.uu.nl/~4075897/UXopdr2/

Phone
http://www.students.science.uu.nl/~4075897/UXopdr2Mobile/

Tablet
http://www.students.science.uu.nl/~4075897/UXopdr2Tablet/