

Testing the Reliability of Fitts' Law on Different Devices

Basim Alsalmi

Computer-Human Interaction Research Lab (CHIRL),
Department of Mathematics and Computer Science
Laurentian University, Sudbury, Canada
balsalmi@laurentian.ca

Ratvinder Grewal

Computer-Human Interaction Research Lab (CHIRL),
Department of Mathematics and Computer Science
Laurentian University, Sudbury, Canada
rgrewal@laurentian.ca

Abstract – Nowadays, the first sources of information on the Internet are websites with its categories. Some users find certain websites' layout and design not always user friendly. This may lead to poor performance, ambiguity and loss of potential customers. Websites designs' use the principles of user interface (UI) laws and usability guidelines to be consistent and convenient to users. Due to the evolution of technology, society now uses various devices with a wide range of features like screen size, in order to obtain their information. Examples include personal computers (PC), smartphones (SP,) and tablets (TB). Since user interface laws were founded in the 19th and 20th centuries and most were applicable for the PC's websites, we test and investigate the reliability of one of the user interface laws, Fitts' law, on three devices: PC, SP and TB. User interfaces were designed with one task, and the task represented the Fitts' law. Human-Computer Interaction (HCI) techniques, user interface design and evaluation methods were followed to test the reliability of Fitts' law on different devices.

Keywords – Usability, HCI, Fitts' law, User Interfaces, Predictive Model, Interaction Design.

I. INTRODUCTION

The component of an interactive computer system that interacts and communicates with the user is the user interface. Any aspect of the system that the user can see and deal with is considered a design of the user interface. In computer systems, the user interface is the component that is becoming a more focused part of the software. As more users utilize computers, the user interface is becoming the most important part of the process. The user interface is the crucial bottleneck in applying a computer-based system to resolve the problems [1], as computers become more robust and powerful.

Since the design of the user interface comprises all the components that are visible to the user, the design of the interface stretches deep into the entire interactive system design. After the system is built, a proper user interface cannot be simply added to it. The interface needs to be a part of the design process from the beginning [1]. There are numerous advantages of having a proper user interface and can make a fundamental difference in user's performance speed, error rates and training time. However, poor designs can prevent the user from continuing to use the interface. To guide the designers and evaluators, theories and predictive models of performance are developed.

In the past, users only used the personal computer (PC) to access the Internet. Nowadays, there are many devices used to access the Internet, such as smartphones (SP) and tablets (TB). In December of 2013 [2], 6% of the global

population owned a tablet, 20% owned PCs and 22% owned smartphones, according to historical data by BI Intelligence, Business Insider's subscription research service. It also reported that the universal smartphones exploded from 5% of global population in 2009 to 22%. Therefore, in four years there was an increase of approximately 1.3 billion smartphones, which means that for every nine people there were two smartphones. Tablets showed quicker adoption rates than smartphones and achieved in two years what smartphones achieved in four years. This indicates that one in every 17 people owns a tablet [2]. These findings show that people use smartphones and tablets more nowadays to interact and access the Internet.

II. BACKGROUND

Fitts' law has been successfully integrated within empirical tests. However, it has received numerous critical reviews as provided [3] [4]. This is influenced by the development of an array of shortcomings upon close examination through the development of a comparison with other studies. This has influenced the integration of numerous corrections or refinements pertaining to the model. This section incorporates an analysis of these and other proposals aimed at providing a conclusive guide pertaining to the application of the law.

Formulation for Index of Difficulty

Earlier analysis incorporated pertaining to the law identified the existence of regular departure of data points above the regression line pertaining to the integration of less complex tasks ($ID < 3$ bits). This influenced the integration of a new formulation for ID [5] to eliminate the disparity:

$$ID = \log_2(A/W + 0.5) \quad (1)$$

The majority of researchers including Fitts identified the existence of an enhanced fit through the utilization of Equation 1. The Welford formulation was utilized by Card [6] [7]. This has influenced the adoption of the Welford formulation by the majority of HCI researchers [8] [9].

Several researchers espouse that in the development of the model by Fitts, the theory incorporated immense differences from Shannon's original work pertaining to the information theory [10] [11]. This influenced the development of a theoretical sound formulation for the index of task as illustrated below:

$$ID = \log_2(A/W + 1) \quad (2)$$

In relation to the terms of *MT*, the prediction model is as illustrated below:

$$T = RT + MT = a + b \log_2(A/W + 1) \quad (3)$$

The time *T* is equal to the reaction time *RT* plus the movement time *MT*. Equation 3, which is also referred to as Shannon formulation provides several benefits including:

- The provision of an enhanced fit pertaining to observations.
- Incorporates the policies stipulated by the information theorem pertaining to Fitts' law.
- Influences the provision of positive results regarding the index of task difficulty.

Extension to Two Dimensions

HCI researchers integrating Fitts' law utilize target selection tasks through a two-dimensional CRT display. This necessitates the integration of a conclusive review pertaining to the shape of the target and the angle of approach upon the application of the model. In the event that the targets are circles or squares, one-dimensional constraints remain in existence. However, in the event that the targets are rectangles such as words, the situation is considered more complex. This study views the amplitude as the distance to the center of the target, which provides a vague definition of target width. This is demonstrated in Figure 1.

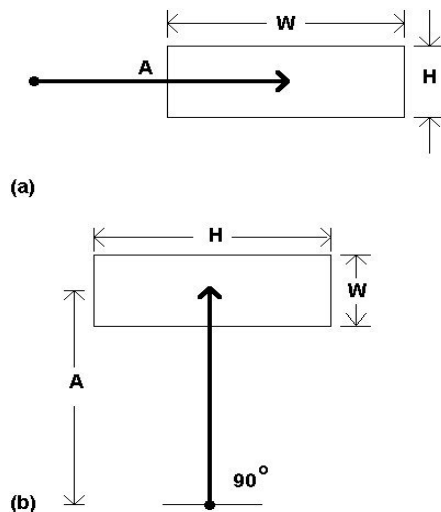


Fig. 1. Fitts' law in two dimensions. The roles of width and height reverse as the approach angle changes from 0° to 90°

The integration of the model within 2D tasks raises questions pertaining to the identification of the target width. This is catered to through the utilization of the horizontal extent of the target, which is referred to as the "STATUS QUO" model for target width. However, this model provides low and at times negative estimates pertaining to task difficulty. This is identified in the case of a short and wide target for instance a word or series of words, which is analyzed through the review of the above or below at close range. Gillan [9] utilized Fitts' law

within a target selection task through the integration of a variety of characters as targets while incorporating varied approaches pertaining to the approach angle and approach distance. This was illustrated by a 26-character (6 cm) target was approached diagonally from a distance of 2 cm. The process integrated the Welford's formulation, which identified the index of task difficulty at $ID = \log_2(A/W + 0.5) = \log_2(2/6 + 0.5) = -2.6$ bits. This was also illustrated by Card [34].

One consequence of the task difficulty extending to the left of $ID = 0$ bits is that it is rendered a positive as it enlarges, which develops under linear regression. This is influenced by the corresponding of conditions with $ID = 0$ bits or less to occurring conditions within the experiment. These tasks incorporate a positive effect the majority of time. However, this may be altered through the integration of two approaches. The first approach involves the utilization of the Shannon formulation for *ID*, which influences the identification of an increase within the rating to $\log_2(2/6 + 1) = +0.42$ bits. The second approach involves the substitution for *W* a measure with one that is in-line with the 2D nature required by the task. Through an analysis of Figure 2, the existent 1D challenge identified pertaining to the constraint with the model is enhanced through the integration of a measurement of *W* along the approach axis.

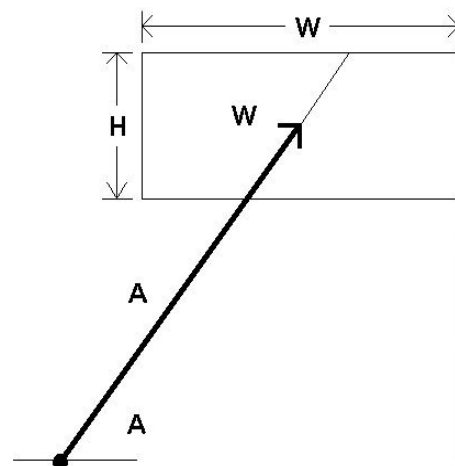


Fig. 2. The smaller of *W* or *H*

In addition the target width may also be substituted with the small capacity of *W* or *H*. This identifies that the smaller of the two dimensions identifies the existence of high levels of accuracy required by the task. This is referred to as the "SMALLER-OF" model, which is identified less complex as it may be incorporated within *A*, *W*, and *H*. The SMALLER-OF model influences rectangular targets. The SMALLER-OF model is more effective than the STATUS QUO model [12]. This influences the recommendation of Fitts' law within two-dimensional tasks [13].

III. METHODOLOGY

In this section, we will discuss the equipment that was used in our study as well as the participants that were

recruited. We will also illustrate the methodology that was utilized for the experiment and the design of user interfaces.

Equipment

It was essential that the equipment used for both conducting an experiment and gathering the data needed to be completely provided. The equipment included a laboratory room, computer, smartphone and tablet.

Computer

A computer with 2 GB of RAM was used to conduct the experiment. The computer had 2 GB of free hard disk space and the screen size of the monitor was 19 inches. A regular keyboard and a regular laser mouse were used in the experiment.

Smartphone

The Samsung Galaxy Note 3 was used in the experiment, which was equipped with a 5.7 inch FHD S-AMOLED display. The resolution of the screen was 1080 x 1920 pixels. It had a suitable processor in regard to mobile devices with a quad-core 2.3 GHz. It came with internal memory of 32 GB and 3 GB of RAM. The 4.3 Jelly Bean Android operating system was used with Touch Wiz user interface and software. The picture of the Samsung Galaxy Note 3 can be seen in Figure 3.



Fig. 3. Samsung Galaxy Note 3

Tablet

The Samsung Galaxy Tab 2 10.1 was used in the experiment. It came with a 10.1 inch PLS TFT capacitive touchscreen display. The resolution of the screen was 800 x 1280 pixels. The processor used was a dual-core that operated at 1 GHz. Samsung has customized the interface with its Touch Wiz UX software. It came with an internal memory of 32 GB and 1 GB of RAM. The operating system was Android 4.0.3 (Ice Cream Sandwich) operating system. The picture of the Samsung Galaxy Tab 2 10.1 can be seen in Figure 4.



Fig. 4. Samsung Galaxy Tab 2 10.1

It is noteworthy that the screen sizes that we dealt with were completely different from each other. It can be said that the PC screen size is approximately double the TB screen size. Also, the size of the TB screen is almost twice the SP screen size. This is shown in Figure 5.

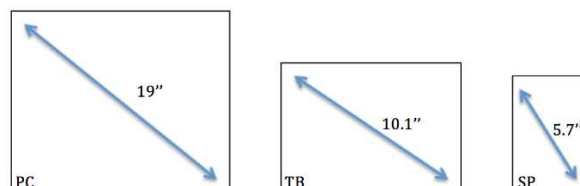


Fig. 5. Display sizes for the three devices

Stopwatch

The Oslo Silver 2.0 twin chronograph and countdown timer was used in the experiment, and only in this experiment. This stopwatch was used to achieve certain role in our task that the participants have completed, which was timing the reaction time. Figure 6 depicts the utilized stopwatch.



Fig. 6. Stopwatch was used in the experiment

Participants

The experiment was conducted using students from Laurentian University. Participants could have been anyone who uses the Internet on a regular basis. The participant was required to perform the task on each device. This information was taken into account to predict how many participants we might need for the experiment and for the sake of obtaining accurate outcomes. The permutation and combination concepts were used to achieve that. Since there are three devices: 3! ($3 \times 2 \times 1 = 6$), the concept was applied to the order of the devices. So,

we had made the experiment for six orders, which are as follows:

- Order 1: PC SP TB
- Order 2: PC TB SP
- Order 3: SP PC TB
- Order 4: SP TB PC
- Order 5: TB PC SP
- Order 6: TB SP PC

The experiment was conducted with 37 participants, consisting of 31 males and 6 females from Laurentian University. The participants used in the study were in different programs, such as Chemical Engineering, Commerce, Computer Science, Economics, and Engineering. There were 21 participants from the Computer Science program and 16 participants from other programs. Also, there were 24 graduate participants and 13 undergraduate participants. There were also 27 participants who were 27 years old or less and 9 participants who were older than 27 years old.

The 37 participants were divided into six groups. Each group consisted of 6 participants, except for group 6, which had 7 participants. The reason for this division was so each group could participate in one of the six orders from the counter-balance principle. The participants were notified that there was a timer working in the background to measure the duration of the experiment to collect data.

User Interfaces Design for the Task: Press on Clickme button then follow the numbers

This task represented the main idea of Fitts' law that is a mathematical model that predicts how long it will take a user to point at a button. It takes into consideration the starting point of the cursor in relation to the button as well as the size of the button and its location. The farther the user needs to move the mouse to reach a button, the more effort this will require from the user. The smaller the button, the more difficult it will be to click. In accordance with this, Task 1 was constructed following the Fitts' law theory. Therefore, many of the interfaces for the Task were designed as shown in Figure 7.

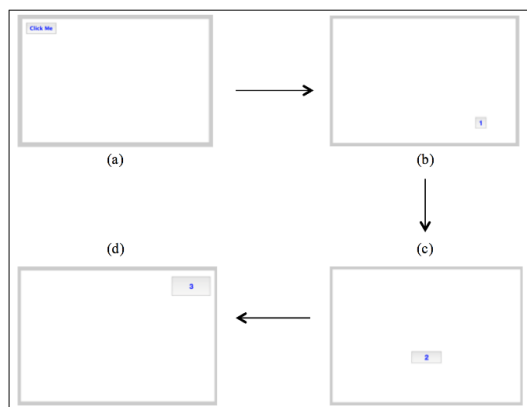


Fig. 7. The interfaces for the Task

When the participant pressed the Clickme button, another button (button "1") of a different size and in a different location appeared. This is shown in Figure 7 (a), (b). The participant was required to click this button. The time taken for the participant to travel from Clickme button to button "1" was recorded in the database for analysis.

After the participant clicked on button "1", another button (button "2") of a different size and in a different location appeared in the middle of the page. The participant was required to click on this button to proceed on in the experiment. The button "2" was similar in size to the Clickme button, but the distance was different between the Clickme button and button "1," and button "1" and button "2," the latter being shorter. This can be seen in Figure 7 (c).

The time taken for the participant to click on button "2" was recorded in the database. After the participant clicked on button "2," the third button (button "3") of a different size and in a different location appeared at the top right corner of the page. It also needed to be clicked on in order to continue the experiment. Button "3" was different in the size of all other buttons and the distance was shorter than the distance between the Clickme button and button "1," but was longer than button "1" and button "2." The time taken for the participant to click on button "3" was recorded for analysis. This can be seen in Figure 7 (d).

IV. RESULTS

The findings of the data analysis are presented in this section. ANOVA was used to identify the statistical differences in the mean time taken for the participants to perform Task 1 on the three devices. The mean differences were tested by one-way analysis of variance with an alpha (P) level of 0.05. The SPSS software was used to analyze the findings of our experiment.

Fitts' law predicted that the time taken to complete the Task would be 2.510 seconds. The Task was divided into 3 sub tasks. We analyzed the data for the Task as one whole task and for each sub task. In the Task a number of outliers were found in the data on the three devices:

- PC – 11 outliers found;
- SP – 6 outliers found;
- TB – 6 outliers found.

After removing the outliers, the overall mean time taken to complete the task on the three devices was calculated. This can be seen in Figure 8.

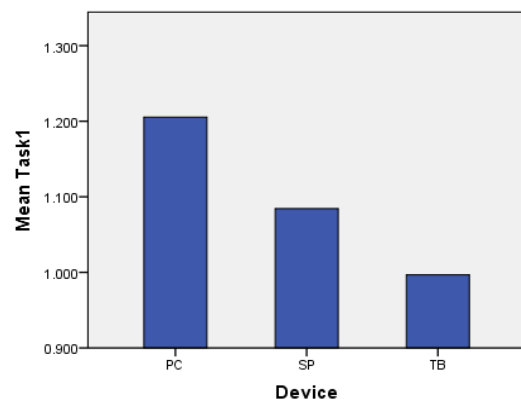


Fig. 8. The mean time taken to perform the Task per device (time in seconds)

The bar chart points out that the participants took less time to perform the task using the TB device, with $\mu = 0.996$ seconds. However, the mean time taken to complete the task using the SP was $\mu = 1.084$ seconds and with the PC was $\mu =$

1.205 seconds. Table I shows the statistics for the Task on the three devices. The second column in the table shows the number of participants in the study. The devices are shown in the first column and the characteristics of the data are shown in other columns.

Table I. Task performed in the experiment (time in seconds)

Descriptives								
Task1								
	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
PC	100	1.20547	.538486	.053849	1.09862	1.31232	.583	3.013
SP	105	1.08444	.354880	.034633	1.01576	1.15312	.653	2.961
TB	105	.99667	.466843	.043607	.91019	1.08314	.551	2.727
Total	310	1.09375	.457999	.026013	1.04257	1.14494	.551	3.013

A significant difference was suggested by the one-way ANOVA results in the mean time taken for the participant to click on buttons, $F(2,307) = 5.512$, $p = 0.004$. The statistics for the time taken using the three devices are illustrated in Table II.

Table II. One-way ANOVA for the time taken to perform the

Task ANOVA					
Task1					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2.247	2	1.123	5.512	.004
Within Groups	62.570	307	.204		
Total	64.817	309			

These are the results of the Task. Next, we will present the results of sub task (a), sub task (b) and sub task (c).

a) The Clickme button to button "1":

This was the first sub task of the Task. Fitts' law predicted that the time taken to complete this sub task would be 3.470 seconds. The sub task was to measure the time taken for a participant to move from the Clickme button to button "1". The data for sub task (a) was analyzed and the number of outliers was found in the data as follows:

- PC – 1 outlier.
- SP – 3 outliers.
- TB – 4 outliers.

After removing the outliers, the overall mean time taken on the three devices was calculated. This can be seen in Figure 9.

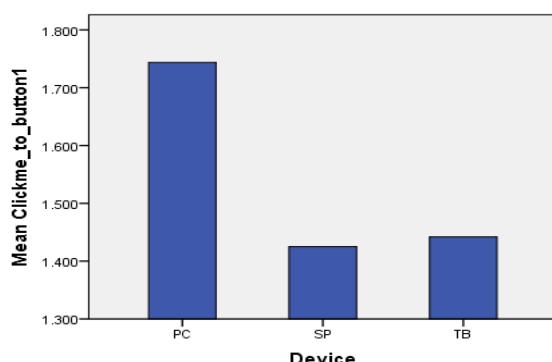


Fig. 9. The mean time taken to perform sub task (a) per device (time in seconds)

The bar chart shows that the participants took less time to perform the sub task (a) using the SP device, with $\mu = 1.424$ seconds, whereas the mean time taken to complete this sub

task using the TB was $\mu = 1.442$ seconds and with the PC was $\mu = 1.743$ seconds. Table III depicts the statistics for sub task (a) on the three devices.

Table III. Sub task (a) performed in the experiment (time in seconds)

Descriptives

Clickme to Button1

	N	Mean	Std. Deviation	Std. Error	95% Confidence Interval for Mean		Minimum	Maximum
					Lower Bound	Upper Bound		
PC	36	1.74375	.541942	.090324	1.56038	1.92712	.803	3.013
SP	34	1.42497	.428418	.073473	1.27549	1.57445	.967	2.961
TB	33	1.44203	.541226	.094215	1.25012	1.63394	.739	2.727
Total	103	1.54185	.523663	.051598	1.43951	1.64420	.739	3.013

The one-way ANOVA suggested that there was a significant difference in the mean time taken to perform the sub task (a), $F(2,100) = 4.397$, $p = 0.015$. The statistics for the time taken using the three devices for sub task (a) are shown in Table IV.

Table IV. One-way ANOVA for the time taken to perform sub task (a)

ANOVA					
Clickme to Button1					
	Sum of Squares	df	Mean Square	F	Sig.
Between Groups	2.261	2	1.130	4.397	.015
Within Groups	25.710	100	.257		
Total	27.971	102			

b) Button "1" to button "2":

This was the second sub task of Task 1. Fitts' law predicted that the time required to complete this sub task was 1.320 seconds. The sub task was to measure the time taken for a participant to move from button "1" and point to button "2". The data for sub task (b) was analyzed to gain the results. The number of outliers in the data when the participants using the PC was 3. Also, no outliers were found when the participants performed the sub task using the SP and the TB. After removing the outliers, the overall mean time taken on the three devices was calculated. This can be seen in Figure 10.

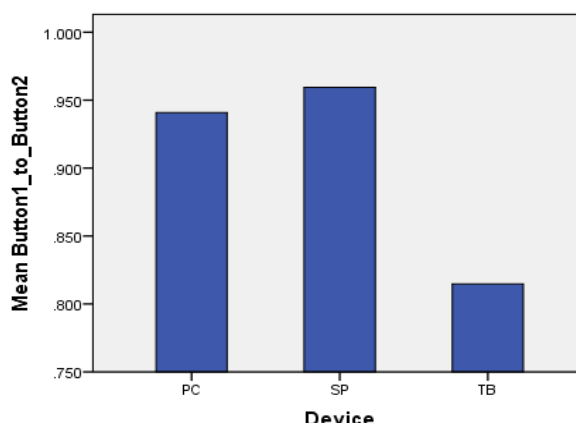


Fig. 10. The mean time taken to perform sub task (b) per device (time in seconds)

The bar chart illustrates that the participants took less time to perform the sub task (b) using the TB device, with $\mu = 0.814$ seconds, whereas the mean time taken to complete this sub task using the SP was $\mu = 0.959$ seconds and with the PC

quickest on the TB. As known, Task 1 was divided into 3 sub tasks. Each sub task will be discussed individually.

a) Clickme button to button “1”:

Pointing at a small size button is a difficult task, especially if the button is located far away from the starting point. The results of sub task (a) revealed that the participants’ performance was better on the SP and the TB, with a slight advantage for the SP. Based on these results, there is a difference in the time taken to point at button “1” on the SP and the TB comparing to the PC.

b) Button “1” to button “2”:

Pointing at a button is an easy task to do when the distance is short. The results for sub task (b) illustrated that the TB was the device on which the participants performed their best. The SP recorded the highest mean time taken to accomplish the sub task (b), with a slight difference from the PC. Based on the results, there is a difference in the time taken to point at button 2 on the TB compared to the other devices.

d) Button “2” to button “3”:

In accordance with the results, pointing at a large button with a short distance is the easiest task to complete. The results of sub task (c) pointed out that the participants’ performances were better on the TB than the PC and the SP. Based on the results, there is a difference in the time taken to point at button 3 on TB comparing to others.

Table IX. Usability VS Modeling

Task#	Usability			Modeling
	PC	SP	TB	
Task 1	1.542	0.904	0.831	2.510
Task 1 (a)	1.744	1.425	1.442	3.470
Task 1 (b)	0.941	0.959	0.815	1.320
Task 1 (c)	0.849	0.880	0.769	2.740

Since our main goal was to test the reliability of Fitts’ law on three different devices, the time taken was predicted for the task using the predictive model. Then, we compared the predicted time by modeling with the mean of the time taken for the participants to accomplish the task. See Table IX.

Even though the predictive models that we utilized were applicable to the PC, the time predicted by the models was different than the time needed by the participants. In fact, the predicted time by the modeling was much longer than the usability time. This means that the predictive model predicted incorrect time for the Task in accordance with Table IX.

This user interface law (Fitts’ Law) was introduced a long time ago, in 1954. In response to this fact, we argue that Fitts’ law is out of date. This is obvious by looking at Table IX where the time predicted by Fitts’ modeling was quite different than the usability time. Also, we suggest that this predictive model need some modifications to be compatible with the modern devices with different screen sizes that we have today.

VI. CONCLUSION

A Task was designed to test one of the user interface laws, which is Fitts’ law. The task was concerned with the time taken to point at buttons on the devices’ screen while taking in consideration the starting point of that movement. The findings of the Task showed that there is a difference in the time taken for the participant to accomplish the task. The participants performed much better to point at buttons on the screen of the TB device than the screen of both the PC and the SP. Moreover, the results indicated that some of the participants using the PC conducted many cases of overshooting to point at the smallest and farthest button when using a mouse to achieve the task.

REFERENCES

- [1] Jacob, R. (2000). "User Interfaces," in *Encyclopedia of Computer Science, Fourth Edition*.
- [2] Heggsetuen, J. (2013). One in every 5 people in the world own a smartphone, One In Every 17 Own A Tablet [CHART]. *Business Insider*. Retrieved 14 January 2015, from <http://www.businessinsider.com/smartphone-and-tablet-penetration-2013-10>
- [3] Meyer, D. E., Smith, J. E. K., Kornblum, S., Abrams, R. A., & Wright, C. E. (1990). Speed-accuracy tradeoffs in aimed movements: Toward a theory of rapid voluntary action. In M. Jeannerod (Ed.), *Attention and performance XIII* (pp. 173-226). Hillsdale, NJ: Erlbaum.
- [4] Welford, A. T. (1968). *Fundamentals of skill*. London: Methuen.E.
- [5] Welford, A. T. (1960). The measurement of sensory-motor performance: Survey and reappraisal of twelve years' progress. *Ergonomics*, 3, 189-230.
- [6] Card, S. K., English, W. K., & Burr, B. J. (1978). Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys for text selection on a CRT. *Ergonomics*, 21, 601-613.
- [7] Card, Stuart K.; Moran, Thomas P.; Newell, A. (1983). *The psychology of Human-Computer Interaction*. Hilldale, London: Lawrence Erlbaum.
- [8] Boritz, J., Booth, K. S., & Cowan, W. B. (1991). Fitts's law studies of directional mouse movement. *Proceedings of Graphics Interface '91*(pp. 216-223). Toronto: CIPS.
- [9] Gillan, D. J., Holden, K., Adam, S., Rudisill, M., & Magee, L. (1990). How does Fitts' law fit pointing and dragging? *Proceedings of the CHI '90 Conference on Human Factors in Computing Systems* (pp. 227-234). New York: ACM.
- [10] MacKenzie, I. S. (1989). A note on the information-theoretic basis for Fitts' law. *Journal of Motor Behavior*, 21, 323-330.
- [11] Shannon, C. E., & Weaver, W. (1949). *The mathematical theory of communication*. Urbana, IL: University of Illinois Press.
- [12] MacKenzie, I. S., & Buxton, W. (in press). Extending Fitts' law to two dimensional tasks. *Proceedings of the CHI '92 Conference on Human Factors in Computing Systems*. New York: ACM.
- [13] MacKenzie, I. S. (1995). Movement time prediction in human-computer interfaces. In R. M. Baecker, W. A. S. Buxton, J. Grudin, & S. Greenberg (Eds.), *Readings in human-computer interaction* (2nd ed.) (pp. 483-493). Los Altos, CA: Kaufmann. [reprint of MacKenzie, 1992].