FFitts Law: Modeling Finger Touch with Fitts' Law

Xiaojun Bi Yang Li Shumin Zhai
Google Inc.
Mountain View, CA, USA
{bxj, liyang, zhai}@google.com

ABSTRACT
Fitts' law has proven to be a strong predictor of pointing performance under a wide range of conditions. However, it has been insufficient in modeling small-target acquisition with finger-touch based input on screens. We propose a *dual-distribution hypothesis* to interpret the distribution of the endpoints in finger touch input. We hypothesize the movement endpoint distribution as a sum of two independent normal distributions. One distribution reflects the relative precision governed by the speed-accuracy tradeoff rule in the human motor system, and the other captures the absolute precision of finger touch independent of the speed-accuracy tradeoff effect. Based on this hypothesis, we derived the *FFitts* model—an expansion of Fitts' law for finger touch input. We present three experiments in 1D target acquisition, 2D target acquisition and touchscreen keyboard typing tasks respectively. The results showed that FFitts law is more accurate than Fitts' law in modeling finger input on touchscreens. At 0.91 or a greater R^2 value, FFitts' index of difficulty is able to account for significantly more variance than conventional Fitts' index of difficulty based on either a nominal target width or an effective target width in all the three experiments.

Author Keywords
Fitts' law; Touchscreen; Finger input

ACM Classification Keywords
H.1.2 [User/Machine System]: Human factors; H.5.2 [User Interface]: Theory and Methods.

INTRODUCTION
Since originally published in 1954, Fitts' law (Eq. 1) [11] has proven to be one of the most robust and successful models of human motor behavior. In HCI, Fitts' Law is typically defined as:

$$T = a + b\log_2\left(\frac{A}{W} + 1\right), \quad (1)$$

where T is the average time taken to complete the movement, A is the distance from the starting point to the center of the target, W is the width of the target, a and b are constants reflecting the efficiency of the pointing system.

Because of its strong predictive power, Fitts' law has served as one of the quantitative foundations for human-computer interaction research and design. It has been used as a theoretical framework for computer input device evaluation [6, 18], a tool for optimizing new interfaces [5, 16], a predictive element in complex gesture recognition algorithms [26], as well as a logical basis for modeling more complex HCI tasks [1].

Dating back from Fitts' original studies [11], target acquisition tasks were typically carried out with a stylus or a cursor that is much smaller than the targets. As finger touch on the popular smart phones and tablets emerges as one of the main input modalities today—the post-PC computing era—examining Fitts' law for finger touch has been attracting attention from HCI researchers [8, 20]. A critical challenge in applying Fitts' law to finger input is that finger input is imprecise, especially relative to smaller-sized targets [8, 20, 14, 13], due to the obvious and well-known "Fat Finger" problem. Previous research showed that Fitts' law's predictive power dropped when targets were small [7, 22]. Our experiments presented later in this paper confirmed such degradation of the conventional forms of Fitts' law for small target acquisition using finger input as well.

To accurately model finger input, we propose a *dual-distribution hypothesis* to interpret the distribution of endpoints of finger input. We hypothesize that the endpoint distribution is a sum of two independent normal distributions. One reflects the relative touch precision governed by the speed-accuracy tradeoff in the human motor system, and the other reflects the absolute precision of finger touch independent of the speed-accuracy tradeoff effect.

Based on this hypothesis, we derive the FFitts model—an expansion and also a refinement of Fitts' law for finger touch input (Eq 2). Our study results show that the FFitts model is strong in predicting finger touch input performance, and it outperforms conventional Fitts' law with either a nominal target width (Eq. 1, hereafter referred to as ID_n model) or an "effective target" width (Eq. 3, hereafter referred to as ID_e model) in both 1D and 2D Fitts' aimed movement tasks, as well as in a text entry task in which the Digraph-Fitts model [16] was used to predict text entry speed.

The contributions of this paper are two-fold:

1) We propose the dual-distribution hypothesis to interpret the distribution of endpoints for finger input.

2) We derive the FFitts law (Eq. 2) model based on the dual-distribution hypothesis:

$$T = a + b\log_2\left(\frac{A}{\sqrt{2\pi e(\sigma^2-\sigma_a^2)}}+1\right) \quad (2)$$

where σ is the standard deviation of the touch points, and σ_a reflects the absolute precision of the input finger, which is independent of the task.

To our knowledge this is the first time Fitts' law has been systematically and successfully extended to finger input on phone-sized touchscreens. In three experiments we show that this model is able to better predict finger touch performance than the conventional forms of Fitts' law.

RELATED WORK

Effective Width Adjustment Method

Fitts' law in its original form predicts human (performer) movement time from the nominal task parameters of target distance (A) and target width (W). The logarithm of the ratio A/W, measured in bits, was viewed as the task's index of difficulty. It was realized that the performer may over- or under-utilize the target size W. In other words the performer's actual pointing precision could be different from the nominal task specification [9, 18, 21]. The most common way of compensating for this discrepancy is to replace the nominal target width W with the so-called effective width, $W_e = \sqrt{2\pi e}\sigma$, hence:

$$T = a + b\log_2\left(\frac{A}{\sqrt{2\pi e}\sigma}+1\right) \quad (3)$$

The justification for the use of W_e is commonly traced to Welford [21], which in turns attributes it to Crossman [9]. Crossman's reasoning of W_e relies on an information-theoretic metaphor. $\log_2 W$ was viewed as the entropy of the endpoint distribution. Since endpoints are observed to be normally distributed about the center of the target, the theoretically correct expression for endpoint entropy $H(o)$, is $H(o) = \log_2\sqrt{2\pi e}\sigma$ [9,18, 21].

Although this information-theoretic foundation is only metaphorical without stronger or more rigorous basis, adjusting effective width based on $\sqrt{2\pi e}\sigma$ has been advocated by many researchers of Fitts' Law. For example, MacKenzie [18] suggested that "this adjustment lies at the very heart of the information-theoretic metaphor that movement amplitude area analogous to 'signals' and end-point variability (via target width) is analogous to 'noise'." (section 3.4 Effective Target Width, paragraph 2, p 106).

Recently, Zhai et al. [25] empirically investigated the effect of using effective width vs. nominal width. Their work showed (although not completely) that ID_e partially compensated for subjective accuracy choice and reduced the discrepancy of a and b estimates between different experimental conditions. The R^2 value of T vs. ID_e regression across different operating biases was higher than the R^2 value of T vs. ID_n regression.

Given the justification from information-theoretic metaphor and empirical foundation, adjusting W based on $\sqrt{2\pi e}\sigma$ has been widely adopted if the observed error rates deviate from 4%. In this paper, we are particularly interested in whether ID_e would also compensate for a finger's imprecision and compare the ID_e model with the proposed FFitts index of difficulty ID_f.

Fitts' Law and Finger Input

There has been an increasing interest in understanding the "Fat Finger" problem, and examining Fitts' Law for finger input.

As a key input modality for touchscreens, touch input has been extensively studied by many researchers. Holz and Baudisch's research [13] showed that the offsets of touch point locations from the intended point were affected by the angles between the finger and the touch surface (i.e., pitch, roll and yaw). In the following studies [14], they discovered that users relied on the visual features of fingers such as finger outlines and nail outlines for placing the touch points. As touchscreen hardware usually reported the centroid of the contact area between the finger and the touchscreen as the touch point, the registered position could be very different from the perceived touch point.

Cockburn et al. [8] recently compared finger input, stylus and mouse in tap, drag and radial pointing tasks. The results showed that the completion time of finger tap and drag strongly conformed to the ID_n model, though the error rate was high (around 12%) when W = 5 mm. W values in their study varied in a wide range (W = 5, 12.5, and 20 mm). They did not particularly investigate small target acquisition tasks.

Sasangohar et al. [20] conducted a Fitts' Reciprocal Tapping task to evaluate mouse and touch input on a tabletop display. Their study also showed very high error rates when targets were small: the error rates were above 20% with targets in which W = 5mm. They did not report the regression results for Fitts' law.

Lee and Zhai [15] studied soft button finger tapping performance on smartphones. It is interesting to note that Fitts' law in its traditional form clearly did not work well for their tasks. When the target was smaller, finger touch performance degraded much faster than Fitts' law would have predicted.

Fitts' Law in Small Target Acquisition Tasks

Although Fitts' law has proven to be a strong and robust model under a wide range of conditions, its prediction power drops in small-target acquisition tasks. Welford et al. [22] observed a departure of ID_e model prediction from the actual completion time in a pencil on paper tapping task. They proposed using $W_e - c$ instead of $W_e (= \sqrt{2\pi e}\sigma)$, in Eq. 3, where c was an experimentally determined constant attributed to hand tremor. The modified version gave a good fit to the observed results.

Chapuis and Dragicevic [7] also observed a clear departure from Fitts' law for small target acquisition using a mouse. Their studies confirmed the existence of a small-scale effect that violates Fitts' law, and the causes are both visual and motor. They empirically demonstrated that the c constant adjustment as originally proposed by Welford in passing, was effective in their dataset.

In sum, there is sufficient evidence in the literature to show Fitts' law's degradation in small target acquisition tasks and particularly so when the implement of the acquisition is a "fat finger".

MODELING FINGER TOUCH WITH FITTS' LAW

Dual-Distribution Hypothesis

Fitts' law in essence reveals a speed-accuracy tradeoff rule in human control performance. The less precisely the task is (e.g., acquiring a wider target over a shorter distance), the faster it is to accomplish the task, and vice versa. As revealed in Eq. 1, the task precision is specified by $(D + W)/W$ and speed is measured in the movement time T.

If the performer does not comply with the precision specified by the nominal task parameters, the "effective width" adjustment has been suggested in lieu of W [9, 18, 21]. An underlying assumption behind the "effective width" adjustment is that the variability of endpoints is solely determined by the speed-accuracy tradeoff rule in the human motor system. Therefore, the whole variability in the endpoints (σ) is taken into account when estimating the "effective width" ($W_e = \sqrt{2\pi e}\sigma$).

In finger input, this assumption faces challenges. Obviously a finger per se is less precise than a mouse pointer or a stylus. Variability in endpoints was observed no matter how quickly/slowly a user performed the task. For example, Holz and Baudisch's studies [13, 14] showed that even when users were instructed to take as much time as they wanted to acquire a target on a touchscreen, there was still a large amount of variability in endpoints.

These observations indicated that a portion of the variability in endpoints is independent of the performer's desire to follow the specified precision and cannot be controlled by a speed-accuracy tradeoff. This portion of variability reflects the absolute precision of the finger input. In other words, the observed variability in the endpoints may originate from two sources: the relative precision governed by the speed-accuracy tradeoff of human motor systems, and the absolute precision uncertainty of the finger per se.

More formally, we propose a dual normal distribution hypothesis to interpret the distribution of endpoints for finger input. Assuming that the location of the endpoints is a random variable X following a normal distribution $X \sim N(\mu,\ \sigma^2)$, we hypothesize that X is the sum of two independent, normally distributed random variables, $X_r \sim N(\mu_r,\ \sigma_r^2)$ and $X_a \sim N(\mu_a,\ \sigma_a^2)$.

X_r, which can be viewed as a relative component, depends on the desired precision of hitting the target and the associated speed of the action. It is controlled by the speed-accuracy tradeoff of the performer and reflects a precision relative to the movement amplitude. The faster the performer, the wider dispersion X_r has.

X_a, which can be viewed as an absolute component, is independent of the performer's desire of following the specified task precision, and cannot be controlled by the speed-accuracy tradeoff. It reflects the absolute precision of a motor system that includes the implement (e.g., the finger, or a stylus) and the internal human motor control system.

In other words, assuming that X_r and X_a are independent, the relation of X, X_r and X_a can be written as:

$$X = X_r + X_a \sim N(\mu_r + \mu_a, \sigma_r^2 + \sigma_a^2) \quad (4)$$

or

$$\mu = \mu_r + \mu_a \quad (5)$$

$$\sigma^2 = \sigma_r^2 + \sigma_a^2 \quad (6)$$

Note that the coordinate origin of X, X_r and X_a is defined at the center of the target. Figure 1 illustrates the dual distribution hypothesis in 1D Fitts' tasks.

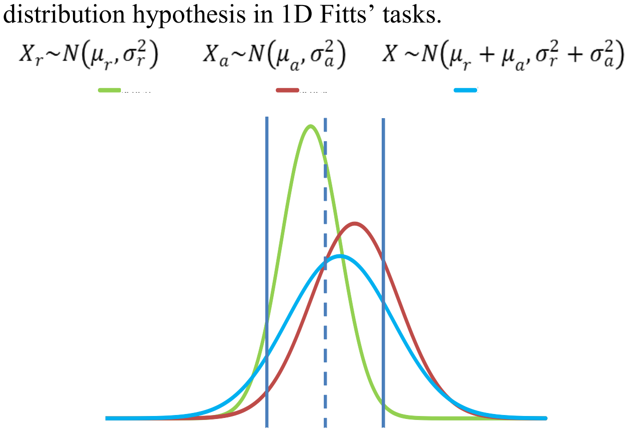


Figure 1. Dual distribution hypothesis in 1D Fitts' tasks. The two solid vertical lines represent the target, and the dashed line is the target center. The green, red and light blue curves show distributions of X_r X_a, and X.

FFitts Law

The dual distribution hypothesis suggests that X_r is the distribution that reflects the variability of endpoints resulted from the speed-accuracy tradeoff effect. It is logically more sound to use σ_r in lieu of σ to express the actual precision performers comply with. Using σ_r in lieu of σ in Eq (3), we have:

$$T = a + b\log_2\left(\frac{A}{\sqrt{2\pi e}\sigma_r} + 1\right) \quad (7)$$

According to Eq (6), σ_r can be obtained as:

$$\sigma_r = \sqrt{\sigma^2 - \sigma_a^2} \tag{8}$$

From Eqs 7 and 8, we have:

$$T = a + b \log_2 \left(\frac{A}{\sqrt{2\pi e (\sigma^2 - \sigma_a^2)}} + 1 \right) \tag{9}$$

We refer Eq. (9) as FFitts Law, a refinement of **F**itts' law for **F**inger input. The index of difficulty of FFitts law is:

$$ID_f = \log_2 \left(\frac{A}{\sqrt{2\pi e (\sigma^2 - \sigma_a^2)}} + 1 \right) \tag{10}$$

FFitts law is also referred as the ID_f model in the remainder of this paper.

In Eqs. 8, 9 and 10, σ is the standard deviation of the distribution of endpoints, which can be measured directly. σ_a reflects the absolute precision of the input finger. It may vary with individual's finger size or motor impairment (e.g., tremor, or lack of).

σ_a can be measured in a *finger calibration task*, in which participants repeatedly acquire a target without specifying an amplitude on the screen. Because such tasks do not involve human motor systems traveling from one place on the screen to another, the speed-accuracy tradeoff rule has a negligible effect. To correctly measure the absolute precision of the input finger, the targets must be sufficiently small yet legible. Accomplishing such a task should require the highest possible precision the user could achieve. A break should be enforced between trials so that the user has sufficient time to attempt to accurately acquire the target. Note that estimating σ_a in this *finger input calibration task* means that σ_a in FFitts law (Eq. 9) is not an additional arbitrary or free-floating parameter in FFitts' law—adding σ_a does not add additional degrees of freedom for the model to fit the data during linear regression analysis.

In our implementation of calibration tasks, the targets were 2.4 mm wide bars (1D tasks) and 2.4 mm-diameter circles (2D tasks). A 1000 ms break was enforced between trials. Alternatively, single pixel wide lines and cross hairs could be used in lieu of bars and circles.

If σ_a is negligible relative to σ, $\sqrt{\sigma^2 - \sigma_a^2} \approx \sigma$. ID_f (Eq. 10) becomes an approximation of the ID_e model (Eq. 3). It could happen if the input device is highly accurate (e.g., a stylus with a thin tip), or the target width is sufficiently large and performers fully utilize it. In the former, $\sigma_a \approx 0$, while in the latter, $\sigma \gg \sigma_a$.

We expect FFitts law would have stronger prediction power than the ID_n and ID_e models, especially than the ID_e model, because σ_r is more accurate than σ in reflecting the actual precision human motor system in compliance with.

To validate the dual distribution hypothesis and evaluate the effectiveness of FFitts law, we conducted the following three experiments.

EXPT 1. ONE-DIMENSIONAL FITTS' TASKS

We first examine the ID_f model in 1D Fitts' tasks (i.e., acquiring bar-shaped targets).

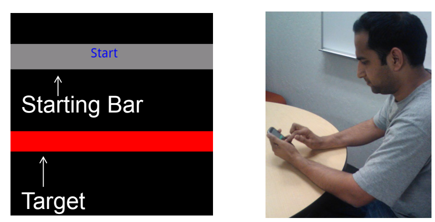
Participants and Apparatus

We recruited 12 participants (3 females) between the ages of 18 and 45. All of them were right-handed and used touchscreen devices (e.g., smartphones) several times a day. The experiment was conducted on an HTC NEXUS smart phone running Android OS. The capacitive touch screen was 48mm wide and 80mm high with a resolution of 480 ×800 pixels. When a finger touched the screen, the approximate centroid of contact area between the finger and the screen was reported as the touch point.

Design

Each participant performed both 1D Fitts' tasks and finger input calibration tasks. The orders of these two tasks were counterbalanced among participants.

1D Fitts' Task. This was a typical "discrete" 1D Fitts' task. At the beginning of each trial, the smart phone played a beep sound and displayed a 6 mm wide horizontal grey bar as the starting bar, and a red horizontal target bar with varying width (Figure 2). Upon successfully selecting the starting bar, the starting bar turned green and the smartphone played a starting beep sound, indicating the start of a trial. The starting bar appeared at a random position for each trial.



(a) (b)

Figure 2. (a) 1D Fitts' Tasks (b) Experimental Setup

The study included 6 $A \times W$ conditions, with 2 levels of A (20, 30 mm) crossed with 3 levels of W (2.4, 4.8, 7.2 mm), yielding ID_n ranging from 1.58 – 3.17 bits. A was measured from the center of the starting bar to the center of the target bar, and W was the width of the target bar. Each $A \times W$ combination included 16 trials. The target widths were chosen to reflect the sizes of common UI elements on smartphones. For example, a hyperlink on a webpage is approximately 2.5 mm wide and a key on a smart phone soft keyboard is around 4 mm wide. The orders of trials were randomized for each participant.

Finger Calibration Task. At the beginning of each trial, the smart phone played a beep sound and displayed a 2.4 mm wide red target bar across the screen. Participants were instructed to acquire the target with the index finger once the trial started. They then lifted their fingers off the screen and rested them comfortably. Each participant performed

16 trials that were divided into 2 blocks. The locations of targets were randomized. The interval between trials was 1000ms.

In both tasks, participants were instructed to acquire the target as quickly and accurately as possible. They performed tasks in hold-and-tap postures: holding the touchscreen device using the non-dominant hand and acquiring the targets with the index finger of the dominant hand. It is one of the most common postures of using a touchscreen device [3].

The target turned yellow and the smartphone played a success beep sound if the touch point hit the inside of the target. Otherwise the target turned blue and played an unpleasant failure sound. After each trial, the number of successful trials M and the number of total trials thus far N in the current block was displayed on the top right corner of the screen in the format of M/N.

The completion time was the elapsed time between the starting beep sound and the participant finished the acquisition task (i.e., the finger was lifted up). The error rate was the percentage of trials failed.

We label touch points that were more than 20 mm away from the target center as "outliers" and removed them from the recorded data. Six trials were removed, consisting of a small percentage of the total number of trials (3,456).

Touch Points for Selection

A user's touch action can generate a series of touch points. We can view a touch action in three stages: land-on, on-screen, and take-off. Land-on refers to the moment when the finger first contacts the screen; on-screen is the state during which the finger remains in contact with the screen; take-off refers to the moment when the finger is lifted off the screen. There are different ways to determine the selection point based on these different stages, which may give different error rates.

In the experiment the mean error rates were 17.8% (Land-On), 16.8% (Take-Off), and 17.02% (Centroid of touch points). ANOVA did not show a significant main effect of acquisition point type on error rate ($F_{2,22} = 2.08$, $p = 0.15$), indicating that touch point type had a minor impact on touch performance. In what follows we use take-off position as the default touch point in the subsequent sections without additional comments.

Results

Table 1 shows the error rate per $A \times W$. When W = 2.4 and 4.8, the error rates were substantially higher than 4%, indicating that participants did not comply with the nominal task precision in these conditions.

Participants tended to over-utilize the target region on small targets. When $W = 2.4$, the Effective Width (W_e) which indicates the actual spread of touch points was more than twice the nominal target width (W).

A	W	σ	σ_r	$W_e = \sqrt{2\pi e}\sigma$	Error Rate
20	2.4	1.21	0.76	5.0	29%
30	2.4	1.33	0.94	5.5	38%
20	4.8	1.52	1.19	6.3	14%
30	4.8	1.52	1.20	6.3	11%
20	7.2	1.66	1.37	6.9	3%
30	7.2	1.78	1.52	7.4	6%
σ_a	0.94 mm				

Table 1. σ, σ_r , σ_a ,W_e (mm) and Error Rate per $A \times W$.

Figure 3 shows σ and σ_r per W_e. As indicated by the dual distribution hypothesis, σ overestimates σ_r in all conditions. The difference between σ and σ_r is especially big when W_e is small. According to the Eq 10, the ID_f model becomes an approximation of the ID_e model if the effective width is large. Figure 3 visually demonstrates it: as W_e becomes bigger, σ and σ_r tend to converge.

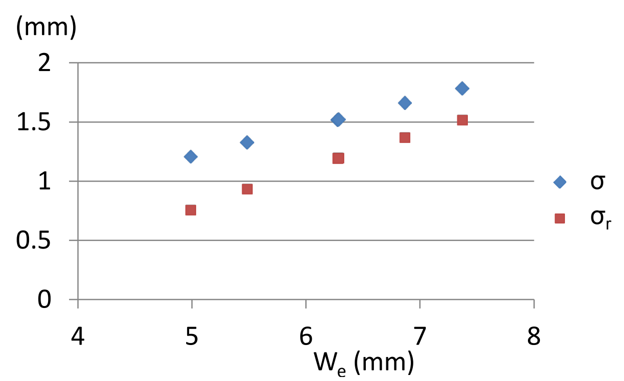


Figure 3. σ and σ_r per W_e in 1D Fitts' tasks

FFitts Law vs. Fitts' Law

We applied ID_f , ID_e and ID_n models separately to the collected data. For the ID_f model, σ_a was the standard deviation of the touch point distribution measured in the *finger calibration tasks* (Table 1).

Figure 4 shows detailed results from the regression tests. As shown, the ID_f model shows a stronger fit than the ID_e, and a slight improvement over the ID_n model. 96% of variance in T could be accounted for by the change of ID_f . In particularly, the ID_f model showed a marked improvement over the ID_e model: R^2 value increased by 11%, from 0.86 to 0.96. The ID_e model appears to be a poor adjustment for target utilization rate in the case of finger pointing.

To gain deeper understanding of the difference between different models, we compared ID values across $A \times W$ conditions (Table 2).

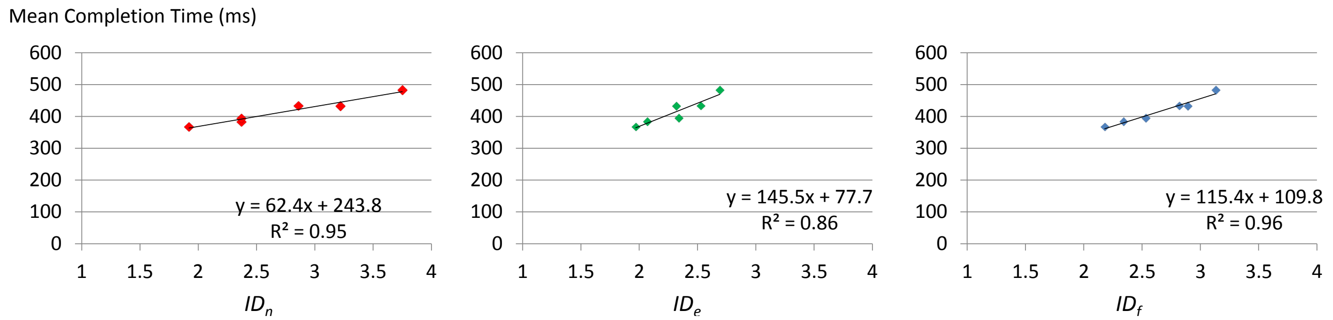


Figure 4. T vs. ID_n (left), T vs. ID_e (middle), and T vs. ID_f (right) for 1D Fitts' Tasks.

A (mm)	W (mm)	ID_n	$W_e = \sqrt{2\pi e}\sigma$	ID_e	$\sqrt{2\pi e} \times \sqrt{\sigma^2 - \sigma_a^2}$	ID_f	Time (ms)
20	2.4	**3.22**	5.0	**2.32**	3.1	**2.89**	432
30	2.4	**3.75**	5.5	**2.69**	3.9	**3.13**	483
20	4.8	**2.37**	6.3	**2.07**	4.9	**2.34**	383
30	4.8	**2.86**	6.3	**2.53**	4.9	**2.82**	433
20	7.2	**1.92**	6.9	**1.97**	5.7	**2.18**	367
30	7.2	**2.37**	7.4	**2.34**	6.3	**2.53**	394

Table 2. ID_n, ID_e, and ID_f per $A \times W$ in 1D Fitts' Tasks.

ID_f vs. ID_e. The following observations can be made by comparing ID_f with ID_e models from Table 2.

First, ID_f was higher than ID_e across all the $A \times W$ conditions. The difference between ID_f and ID_e is especially big in small target conditions. From ID_e to ID_f, the ID values increased by 24.5% (A = 20) and 16.3% (A = 30) when the target width was 2.4 mm, while they only increased by 10.6% (A = 20) and 8.1% (A = 30) when the target width was 7.2 mm.

Second, overall there was an expansion of the range of ID values from ID_e to ID_f. The $|max(ID_e) - min(ID_e)|$ is 0.72, while $|max(ID_f) - min(ID_f)|$ is 0.95.

ID_f vs. ID_n. The following observations were made by comparing the ID_f with ID_n models.

First, ID_f was smaller than ID_n in W = 2.4, while it was bigger when W = 7.2 . ID_f was close to ID_n in W = 4.8.

Second, overall there was a shrinkage of the range of ID from ID_n to ID_f. The $|max(ID_n) - min(ID_n)|$ was 1.83, while $|max(ID_f) - min(ID_f)|$ was 0.95.

In summary, the ID_f model showed the strongest fit among the three models. It especially showed a drastic improvement over the ID_e model.

EXPT 2. TWO-DIMENSIONAL FITTS' TASKS

In addition to the 1D situation, we also investigated 2D Fitts' tasks. The experimental design was the same as in the 1D tasks, except that targets were solid circles and W values were diameters. The experimental devices and the participants were the same as those in 1D Fitts' tasks.

Results

We use the bivariate endpoint deviation (SD_{xy}) as σ. Wobbrock et al. [23] showed that using SD_{xy} yields a better model fit than using univariate deviation (SD_x), which ignores the deviation in the orthogonal task dimension. The SD_{xy} of touch points in the *2D finger calibration* tasks was 1.5 mm, which was used as σ_a in the ID_f models (Table 3).

A	W	σ	σ_r	$W_e = \sqrt{2\pi e}\sigma$	Error Rate
20	2.4	1.68	0.76	6.9	66%
30	2.4	1.72	0.84	7.1	65%
20	4.8	1.84	1.07	7.6	25%
30	4.8	1.88	1.14	7.8	27%
20	7.2	1.98	1.29	8.2	7%
30	7.2	1.97	1.27	8.1	7%
σ_a					1.5 mm

Table 3. σ, σ_r, σ_a, W_e (mm) and Error Rate per $A \times W$.

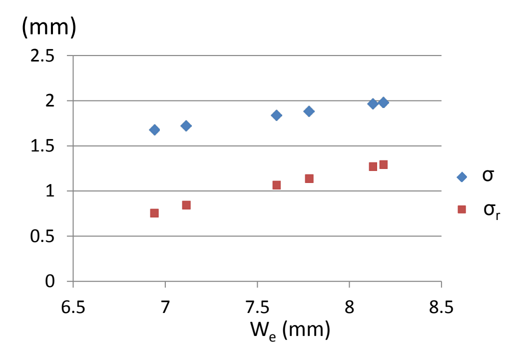


Figure 5. σ and σ_r per W_e in 2D Fitts' tasks

We first investigated the error rates and dispersion of touch points (Table 3). Similar to 1D Fitts' tasks, error rates in W = 2.4 and 4.8 conditions were substantially higher than 4%. Participants also tended to over-utilize the target region as targets become smaller. When W = 2.4, the effective width (W_e) was more than twice of the nominal target width. Similar to the finding in 1D task, Figure 5 shows that σ is greater than σ_r in all conditions, and the difference between σ and σ_r is especially big when W_e is small. Also, as W_e becomes bigger, σ and σ_r tend to converge.

FFitts Law vs. Fitts Law

Figure 6 show regression results for different models. Similar to the findings in the 1D Fitts task, the ID_f model showed the strongest fit. 96% of variance of T could be accounted for by the changes in ID_f. The R^2 value of the ID_f model is 21.5% and 12.9% higher than those of the ID_e and ID_n models. The improvement of R^2 values from ID_e and ID_n to ID_f was greater than that in the 1D experiment.

DISCUSSION OF EXPT.1 AND EXPT.2

FFitts Law is strong in modeling finger input

FFitts law shows strong predictive power for finger input. Approximately 96% of variance in T could be accounted for by the change of ID_f, in both 1D and 2D tasks.

FFitts law is also stronger than both ID_n and ID_e models in predicting finger touch performance, especially in 2D tasks. Using the same data and number of parameters, the R^2 values of the ID_f model were 21.5% and 12.9% higher than the ID_e and ID_n models respectively.

Based on these results, FFitts law, a refinement and expansion of Fitts' law, appears to be a better modeling tool for small target acquisition using finger input.

"Effective Width" method is a poor choice

The study results showed that Fitts' law with "effective width" adjustment had the poorest performance out of the three candidates. ID_e accounted for less than 80% of the time variance in 2D Fitts' task.

The dual-distribution hypothesis provides an explanation. Because the absolute precision (σ_a) of the finger input is not subtracted from σ, σ overestimates the variability of endpoints related to the speed-accuracy trade-off effect.

Reflected in the ID value, ID_e underestimates the ID_f. Our study data showed that ID_e was smaller than ID_f across all the $A \times W$, in both 1D and 2D tasks. The underestimation became relatively big when targets were small. For example, ID_e were 50% smaller than ID_f in W = 2.4 (A = 20, 2D), while ID_e was just 21% smaller than ID_f in W = 7.2 (A = 20, 2D).

EXPT 3. MODELING TEXT ENTRY BEHAVIOR ON A TOUCHSCREEN KEYBOARD

To further evaluate FFitts Law in practice, we applied it to model touch screen keyboard text entry, a critical and popular research topic [4, 5, 16]. Previous research shows that once users reach expert levels, the typing behaviors are largely constrained by the capacity of human motor system to move fingers, not by the visual search. The time to move the tapping device with a single finger from key i to key j for a given distance (D_{ij}) and key width (W_{ij}) follows the Fitts-Digraph model [16]:

$$MT_{ij} = a + b\, log_2(\frac{D_{ij}}{W_{ij}} + 1) \quad (11)$$

To compare the predictive power of different models, we collected users' index finger tapping data on a touchscreen keyboard, and applied ID_f, ID_n, and ID_e models separately according to Eq. (11) to predict MT_{ij}.

Tapping Data Collection

We first collect users' tapping data on a basic touchscreen keyboard (Figure 7) that provided users with only asterisks as feedback when they entered text. The layout of the letter keys was exactly the same as a regular Android keyboard

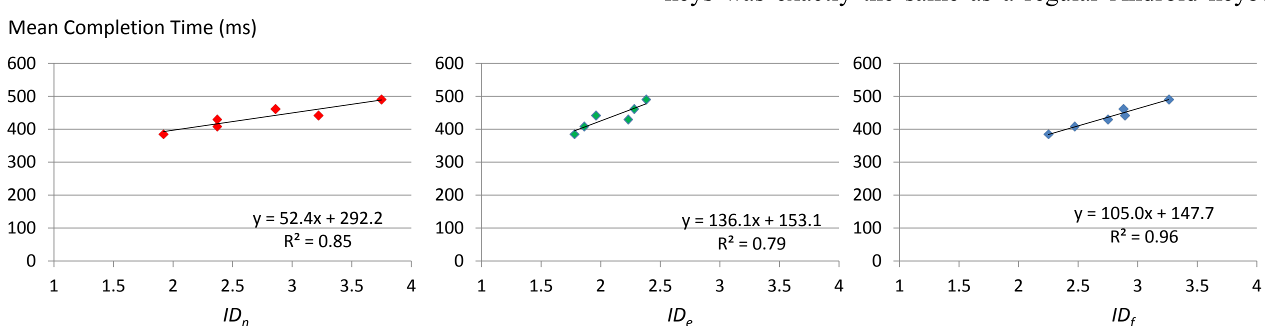


Figure 6. T vs. ID_n(left), T vs. ID_e (middle), and T vs. ID_f (right) for 2D Fitts' Tasks

A (mm)	W (mm)	ID_n	$W_e = \sqrt{2\pi e}\sigma$	ID_e	$\sqrt{2\pi e} \times \sqrt{\sigma^2 - \sigma_a^2}$	ID_f	Time (ms)
20	2.4	**3.22**	6.9	**1.96**	3.1	**2.89**	442
30	2.4	**3.75**	7.1	**2.38**	3.5	**3.26**	491
20	4.8	**2.37**	7.6	**1.86**	4.4	**2.47**	408
30	4.8	**2.86**	7.8	**2.28**	4.7	**2.88**	462
20	7.2	**1.92**	8.2	**1.78**	5.3	**2.25**	385
30	7.2	**2.37**	8.1	**2.23**	5.3	**2.75**	430

Table 4. ID_n, ID_e, and ID_f per $A \times W$ in 2D Fitts' Tasks

on a smartphone. Our purpose was to capture the basic, native and natural text entry behaviors, without being affected by any modern features on smart keyboards, such as auto-correction or auto-completion. It also reflected the expert typing behaviors: typing quickly without checking intermediate results within a word.

Figure 7. A basic keyboard we used to collect typing data.

We recruited 11 participants with an average age of 32. All had experience with text entry on smartphones. Participants entered text with index fingers on a Galaxy Nexus. In the study, each participant repeated a word four times. By repeating a word multiple times in a row, we assumed that the last few repetitions would reflect a user's expert behaviors. This has previously been adopted as an approach to approximate experts' text input behaviors [4, 5].

The whole study included a set of 40 words, which were distributed into 4 blocks. The orders of words within a block were randomized. The study collected 11 (participants) × 4 (blocks) × 10 (words) × 4 (repetitions) = 1,760 words.

Data Processing and Analysis

To align the touch points with the intended keys, we discarded trials in which the number of touch points was different from the number of intended characters. 53 unaligned trials were removed. We labeled touch points as outliers and removed them if the distance between the touch point and the center of the target key was beyond 30 mm.

The input speed increased markedly from repetition #1 to #2, and plateaued after repetition #2. The means of speed (word per min) were 36 at #1, 40 at #2, 39 at #3 and 40 at #4. The results suggest that in the given experimental task users probably reached expert behaviors from repetitions #2 to #4. This finding was similar to the observation from Bi et al.'s work [5]. We used data in repetition #2, #3 and #4 for model evaluation

Note that the input speed was calculated according to the following equation [24]:

$$WPM = \frac{|T-1|}{S} \times 60 \times \frac{1}{5} \qquad (12)$$

where T was the length of the word (number of characters) and S was the elapsed time in seconds from the moment the finger was lifted from the first character and to the moment the finger was lifted from the last character of a given word.

Parameters for Models. Typing a digram on a touchscreen keyboard with one finger is in essence a 2D Fitts' tasks for rectangular targets. Literature has shown different approaches for choosing W for the nominal form of Fitts' law (ID_n model). In this study, we chose the key width (5 mm), as W for ID_n. Key width (5mm) on the keyboard was 30% smaller than key height (7mm). Previous research showed that using *min(W, H)* as W in Fitts' Law yielded a fairly successful fit for 2D Fitts tasks [2, 19].

We used SD_{xy} as σ in the ID_e model. Assuming that σ_a has little variance across tasks, we used 1.5 mm, the value measured in the *finger calibration tasks* in Exp2. as σ_a in ID_f model.

Classifying Digram according to Index of Difficulty. Digram with the same distance between two letters but different start or end characters were classified as the same $A \times W$ condition. In total, the study generated 27 unique nominal ID_n. To ensure that the Fitts' tasks had a sufficient number of sampled touch points, we only picked ID_n which had more than 120 sampled touch points. Nine ID_n were selected, with the ID_n ranging from 1.15 to 2.77. Table 5 shows detailed information of the selected ID_n.

Results

Figure 8 and Table 5 show regression results. The results echoed the findings from the 1D and 2D Fitts' studies: ID_f yields the strongest fit among the three test models. 91% of the variance of completion time could be accounted for by the changes in ID_f. The R^2 values for ID_n, ID_e and ID_f models were 0.75, 0.88, and 0.91 respectively.

In comparison to the conventional Fitts' law formulations based on either ID_n or ID_e, FFitts law also matches better with the keyboard tapping task, which is more complex and less abstract than the 1D and 2D target acquisition tasks. For the text entry task in which the target width, distance and users' behaviors were less controlled than the laboratory Fitts' tasks, FFitts law still showed stronger predictive power than the conventional Fitts' law. Since the Fitts-digrah model, previously developed for stylus typing, has been an important theoretical tool for touch keyboard research, design and optimization [4, 5, 16], improvements found here with the ID_f model constitute an important contribution in their own right.

LIMITATIONS AND FUTURE WORK

Fitts' law serves as a quantitative foundation for many applications on touchscreen devices. For example, besides the sizable body of touchscreen keyboard optimization work [4, 5], Zhai and Kristensson [26] suggested using Fitts' law to refine recognition weights between two channels for word-gesture keyboard recognition. Since FFitts law has stronger predictive power than Fitts' law on smart phones with finger operations, FFitts law should be

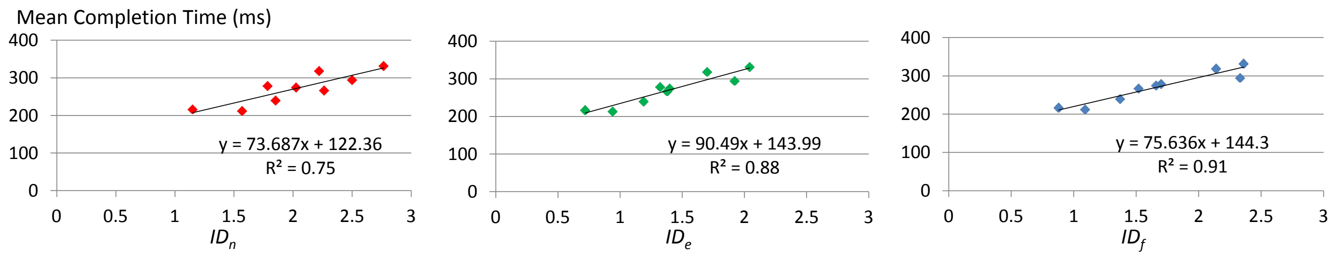


Figure 8. T **vs.** ID_n **(left),** T **vs.** ID_e **(middle), and** T **vs.** ID_f **(right) for word repetition text entry tasks**

Digram examples	A (mm)	W (mm)	ID_n	$W_e = \sqrt{2\pi e}\sigma$	ID_e	$\sqrt{2\pi e} \times \sqrt{\sigma^2 - \sigma_a^2}$	ID_f	Mean (SD) Time (ms)
(e, r), (i, o)	6.19	5	**1.15**	9.57	**0.72**	7.3	**0.89**	217(56)
(e, s), (t, f)	9.98	5	**1.57**	10.74	**0.95**	8.8	**1.1**	213(43)
(o, u), (t, u)	12.38	5	**1.78**	8.28	**1.32**	5.5	**1.7**	260(66)
(e, a), (d, t)	13.24	5	**1.85**	10.35	**1.19**	8.3	**1.38**	240(73)
(c, h), (n, g)	15.57	5	**2.024**	9.52	**1.4**	7.24	**1.66**	275(84)
(i, t), (u, p)	18.58	5	**2.221**	8.24	**1.7**	5.45	**2.14**	319(94)
(b, u), (i, m)	19.26	5	**2.262**	12.03	**1.38**	10.32	**1.52**	267(72)
(a, t), (d, u)	23.65	5	**2.50**	8.51	**1.92**	5.8	**2.33**	295(79)
(a, y), (i, d)	29.41	5	**2.77**	9.44	**2.04**	7.14	**2.36**	332(91)

Table 5. ID_n**,** ID_e**, and** ID_f **per** $A \times W$ **in word repetition text entry tasks.**

adopted in lieu of Fitts' law in these applications The movement amplitudes, target widths and movement time in our three experiments were limited to finger touch interaction tasks on smart phones. We plan to investigate whether the dual distribution hypothesis holds with larger amplitudes and wider targets in the future. A novel part of the hypothesis is that σ_a is introduced to account for the finger's absolute precision, which is an intrinsic characteristic of the finger and not dependent on the movement amplitude. Logically the dual distribution hypothesis and FFitts model are likely to hold for greater amplitudes as well.

On a large touch surface, the target sizes are likely to be larger. The FFitts model suggests that the impact of the absolute finger precision component (σ_a) diminishes very quickly (due to the power function), so the FFitts model will converge with the conventional Fitts' model quickly when the targets are larger. This smooth conversion is a potential strength of the dual distribution hypothesis.

Although the absolute component, σ_a is obtained from a separated calibration task, it can be used to model other Fitts' tasks for the same group of participants. σ_a reflects the absolute precision of finger input, which is independent of other task parameters. Assuming finger size and shape do not vary drastically across users, σ_a could be used across users as an approximation. For example, we used the same σ_a in both Experiments 2 and 3. FFitts law showed strong prediction power in both experiments. In fact, the participants in Experiment 3 were different from those in the calibration task where σ_a was obtained. We suggest that 0.94 and 1.5 mm are used as approximations for σ_a in 1D and 2D Fitts' tasks respectively. This suggestion should also be empirically verified in future work.

CONCLUSION

To accurately model finger input for small target acquisition, we propose the dual distribution hypothesis, and derive the FFitts model, which is an expanded and refined form of Fitts' law. It simplifies to the conventional form of Fitts' law if the absolute precision factor σ_a^2 is negligible in comparison to the target size or the total variance in the task σ^2**.** Three experiments showed that the predictive power of FFitts law is superior to the conventional forms of Fitts' law. Our investigation has led to the following conclusions.

First, FFitts law is a strong model for predicting finger touch performance in small-sized target acquisition tasks. ID_f, the index of difficulty of FFitts law, accounts for more than 91% of the time variance in all three experiments.

Second, FFitts law has stronger predictive power than both the ID_n and ID_e models. FFitts law is especially more accurate than Fitts' law using "effective width". The R^2 values of the ID_f model were 11% (1D tasks) and 21.5% (2D tasks) higher than those of the ID_e model.

Third, neither the ID_n or ID_e model is a strong model for finger input, especially the ID_e model. ID_e accounts for less than 80% of the time variance in 2D Fitts' task. It suggests that the "effective width" adjustment might not be an appropriate choice for finger input.

Fourth, the strong fit of the ID_f model validates the dual-distribution hypothesis, which provides a more logical and reasonable interpretation of the distribution of endpoints than the typical "effective width" interpretation.

In summary, small target acquisition using finger touch input can be accurately modeled by FFitts law:

$$T = a + b\log_2\left(\frac{A}{\sqrt{2\pi e(\sigma^2 - \sigma_a^2)}} + 1\right), \qquad (13)$$

where σ is the standard deviation of touch point distribution, and σ_a is the absolute precision of the input finger.

REFERENCES

1. Accot, J., and Zhai, S. Beyond Fitts' Law: Models for Trajectory-Based HCI Tasks, *ACM CHI,* 295-302.
2. Accot, J. and Zhai, S. (2003). Refining Fitts' law models for bivariate pointing. *ACM CHI,* 193-200.
3. Azenkot, S. and Zhai, S. (2012). Touch Behavior with Different Postures on Soft Smart Phone Keyboards. *Proc. of MobileHCI'12*. 251 - 260.
4. Bi, X., Smith, B., and Zhai, S. (2010) Quasi-qwerty soft keyboard optimization. *ACM CHI*. 283-286.
5. Bi, X., Smith, A. B., and Zhai, S. (2012) Multilingual Touchscreen Keyboard Design and Optimization. *Human-Computer Interaction*, Volume 27, Issue 4, 352 - 382.
6. Card, S.K., English, W.K. and Burr, B.J. (1978) Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys for text selection on a CRT. Ergonomics 21 (8), 601-613.
7. Chapuis, O. and Dragicevic, P. (2011) Effects of motor scale, visual scale, and quantization on small target acquisition difficulty. *ACM Trans. Comput.-Hum. Interact.* 18, 3, Article 13 (August 2011), 32 pages.
8. Cockburn, A., Ahlström, D. and Gutwin, C. (2012) Understanding performance in touch selections: Tap, drag and radial pointing drag with finger, stylus and mouse. *Int. J. Hum.-Comput. Stud.* 70, 3, 218-233.
9. Crossman, E.R.F.W. (1957) The speed and accuracy of simple hand movements. In *The Nature and Acquisition of Industrial Skills*, E.R.F.W. Crossman and W.D. Seymour (eds). Report to the M.R.C. and D.S.I.R. Joint Committee on Individual Efficiency in Industry.
10. Crossman, E.R.F.W. and Goodeve, P.J. (1963/1983) Feedback control of hand-movement and Fitts' law. *Quarterly J. of Exp. Psychology 35A* (2), 251-278.
11. Fitts, P.M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *Journal of Experimental Psychology*, 47, 381-391.
12. Fitts, P.M. and Radford, B.K. (1966) Information capacity of discrete motor responses under different cognitive sets. *J. of Exp. Psychology 71* (4), 475-482.
13. Holz, C. and Baudisch, P. (2010) The Generalized Perceived Input Point Model and How to Double Touch Accuracy by Extracting Fingerprints.*ACM CHI*, 581-590.
14. Holz, C. and Baudisch, P. (2011). Understanding Touch. *ACM CHI,* 2501-2510.
15. Lee, S., and Zhai, S., (2009) The performance of touch screen soft buttons. *ACM CHI*. 309-318.
16. Lewis, J. R., Kennedy, P. J., & LaLomia, M. J. (1999). Development of a Digram-Based Typing Key Layout for Single-Finger/Stylus Input. *Proc. of The Human Factors and Ergonomics Society 43rd Annual Meeting*.
17. MacKenzie, I.S. (1989) A note on the information theoretic basis for Fitts' law. *J. of Motor Behavior 21*, 323-330.
18. MacKenzie, I.S. (1992) Fitts' law as a research and design tool in human-computer interaction. *Human Computer Interaction 7* (1), 91-139.
19. MacKenzie, I.S. and Buxton,W. (1992). Extending Fitts' law to two-dimensional tasks. *ACM CHI*, 219-226.
20. Sasangohar, F., MacKenzie, I.S., Scott, S., (2009) Evaluation of mouse and touch input for a tabletop display using Fitts' reciprocal tapping task. *Annual Meeting of the Human Factors and Ergonomics Society —HFES*, 839–843.
21. Welford, A.T. (1968) *Fundamentals of Skill*. London: Methuen.
22. Welford, A. T., Norris, A. H., and Shock, N. W. (1969) Speed and accuracy of movement and their changes with age. *Acta. Psychol*. 30, 3–15.
23. Wobbrock, J.O., Cutrell, E., Harada S., and MacKenzie, I.S., (2008). An error model for pointing based on Fitts' law. *ACM CHI,* 1613-1622.
24. Wobbrock, J.O. (2007) Measures of text entry performance. In Text Entry Systems: Mobility, Accessibility, Universality, I. S. MacKenzie and K. Tanaka-Ishii (eds.) San Francisco: Morgan Kaufmann, 47-74.
25. Zhai, S., Kong, J. and Ren, X. (2004) Speed-accuracy tradeoff in Fitts' law tasks—on the equivalency of actual and nominal pointing precision. *Int. J. Hum.-Comput. Stud*, 61(6), 823-856.
26. Zhai, S. and Kristensson, P. O. (2013) The word-gesture keyboard: reimagining keyboard interaction. *Commun. ACM* 55, 9 (September 2012), 91-101.