An Error Model for Pointing Based on Fitts' Law

Jacob O. Wobbrock,[1] Edward Cutrell,[2] Susumu Harada[3] and I. Scott MacKenzie[4]

[1]The Information School
[3]Computer Science & Engineering
DUB Group
University of Washington
Seattle, WA 98195 USA
[1]wobbrock@u.washington.edu
[3]harada@cs.washington.edu

[2]Microsoft Research
One Microsoft Way
Redmond, WA 98050 USA
cutrell@microsoft.com

[4]Computer Science & Engineering
York University
Toronto, ON, Canada M3J 1P3
mack@cse.yorku.ca

ABSTRACT
For decades, Fitts' law (1954) has been used to model pointing time in user interfaces. As with any rapid motor act, faster pointing movements result in increased errors. But although prior work has examined accuracy as the "spread of hits," no work has formulated a predictive model for error rates (0-100%) based on Fitts' law parameters. We show that Fitts' law mathematically implies a predictive error rate model, which we derive. We then describe an experiment in which target size, target distance, and movement time are manipulated. Our results show a strong model fit: a regression analysis of observed *vs*. predicted error rates yields a correlation of $R^2 = .959$ for $N = 90$ points. Furthermore, we show that the effect on error rate of target size (W) is greater than that of target distance (A), indicating a departure from Fitts' law, which maintains that W and A contribute proportionally to index of difficulty (ID). Our error model can be used with Fitts' law to estimate and predict error rates along with speeds, providing a framework for unifying this dichotomy.

ACM Categories & Subject Descriptors: H.5.2 [**Information interfaces and presentation**]: User interfaces – *theory and methods*; H.1.2 [**Models and principles**]: User/machine systems – *human factors*.

General Terms: Experimentation, Human Factors, Theory.

Author Keywords: Movement time, pointing time, pointing errors, mousing errors, clicking errors, error rates, speed-accuracy tradeoff, Fitts' law, Schmidt's law, error model.

INTRODUCTION

Even before Newell and Card advocated for a "hardening of the science" of human-computer interaction (HCI) [23], researchers sought quantitative models of human action to explain behavior and inform design. Although there are relatively few such models in HCI, those we do have are highly influential.

CHI 2008, April 5–10, 2008, Florence, Italy.

Perhaps the most influential of these is Fitts' law [8,17]. Since its first application in HCI in 1978 to predict pointing times in a text editor [3,19], Fitts' law has facilitated design innovations [2,10,36], informed aggregate models of computer use [4,13], and been a tool for modeling and evaluation [1,16,18,24,28]. This is no surprise given the law's robustness, ease of use, and the prevalence of pointing in graphical user interfaces.

However, although Fitts' law supports the prediction of speeds, it does not readily support the prediction of errors. In fact, to date, there is no equivalent "error law" that predicts the probability of a user hitting or missing a target using Fitts' law parameters. Although speed-accuracy tradeoffs have been studied (see [12,22,25] for reviews), this work almost universally regards accuracy as the "spread of hits," which is of limited use in predicting error rates in user interfaces. *Post hoc* corrections can be used to normalize differences in speed-accuracy performance among a pool of human subjects [5,17,29,31], but these adjustments lack the predictive power of an error model.

Why predict errors? Error prediction should be as useful as time prediction given the diametric relationship of these two entities: where one increases, the other decreases. Thus, "rounding out" the theory requires a predictive model for errors. Also, if a Fitts-based error model is shown to hold, it contributes to the soundness of the law itself. If it is shown *not* to hold, it motivates a deeper investigation into the assumptions underlying Fitts' law, since, as we show, a Fitts-based error model is mathematically implied.

An error model also has practical applications. For example, it allows us to estimate text entry error rates given different tapping speeds on a stylus keyboard, or to ensure that buttons are big enough in a safety-critical system where speed is crucial. In computer games, as another example, designers may want to predict how many targets a player can hit in a given amount of time.

As we demonstrate, Fitts' law mathematically implies an equation for pointing errors. To our knowledge, this equation has not been derived in the literature. Instead, prior work focuses on motor-control theories accounting for endpoint variability in human movement [6,22,25,26].

Here, we do not seek a motor-control theory. Instead, we acknowledge the influence of Fitts' law and the potential utility of an error model. Accordingly, we derive the Fitts-based error model for pointing and test it in an experiment that manipulates target size, target distance, and movement time. Our results show a strong model fit. A plot of observed (y) *vs*. predicted (x) error rates, where $y = x$ is a "perfect model," yields a regression line with near-zero intercept and near-unity slope: $y = 0.007 + 0.958\,x$ ($R^2 = .959$ for $N = 90$ points). We also show that the effect of target size (W) on error rate is greater than that of target distance (A), indicating a departure from Fitts' law in which W and A contribute equally but inversely to the index of difficulty (ID).

DERIVATION OF AN ERROR MODEL FOR POINTING

Formulated for one dimension, Fitts' law [8] predicts the movement time MT to acquire a target of size W at distance A. Typically, MT is the *dependent* variable in Fitts' law. In MacKenzie's Shannon formulation [15], the law is written

$$MT = a + b \cdot \log_2\left(\frac{A}{W} + 1\right) \tag{1}$$

In Eq. 1, a and b are empirically determined regression coefficients, which vary among users and devices. The log term is known as the *index of difficulty* (ID) and is measured in *bits*. The value $1/b$ is called the *index of performance* (IP) and is measured in *bits/ms*. This has been the measure of *throughput* (TP); an alternative calculation is to use ID / MT [19,29]. Note that the formulation is not concerned with the specific values of A and W, but with their ratio. This gives Fitts' law particular versatility, since TP results can be compared across different experiments. However, comparisons of TP assume that participants perform with similar personal speed-accuracy biases [37]. In fact, owing to its information-theoretic roots, Fitts' law assumes a 4% error rate [17,29,31]. Where a 4% error rate is not observed, a *post hoc* correction attributed to Crossman [5] and discussed by Welford [31] can be made to normalize TP using the effective (observed) target width (W_e) in lieu of the nominal width W. Using W_e allows for mathematically growing or shrinking the effective target so that a 4% error rate *would have been observed*. Although the correction is useful, research has shown that this adjustment is not always wholly corrective [37].

In an equation for predicting errors, the dependent variable gives the probability of an error $P(E)$ and ranges from 0% to 100%. Intuitively, we expect errors to increase as target distance (A) increases, as target size (W) decreases, or as movement time (MT) decreases. Thus, these values are the independent variables in our equation for predicting errors.

Hereafter, we refer to the movement time predicted by Fitts' law as MT, the dependent variable, and the movement time with which someone *actually* moves as MT_e, an independent variable. Owing to tradition [31], we call the latter the "effective movement time." Fitts' law states

$$MT_e = a + b \cdot \log_2\left(\frac{A}{W_e} + 1\right) \tag{2}$$

Now, W_e is the unknown target size coincident with a movement time of MT_e. Solving for W_e,

$$\frac{MT_e - a}{b} = \log_2\left(\frac{A}{W_e} + 1\right) \tag{3}$$

$$2^{\frac{MT_e - a}{b}} - 1 = \frac{A}{W_e} \tag{4}$$

$$W_e = \frac{A}{2^{\frac{MT_e - a}{b}} - 1} \tag{5}$$

In Eq. 5, the target size W_e is proportional to $A / 2^{MT_e}$. Thus, as A increases, W_e increases; as MT_e increases, W_e exponentially decreases. This is the logarithmic speed-accuracy tradeoff captured by Fitts' law.

Prior work shows that the spread of hits in rapid aimed movements forms a Gaussian distribution about the target center [6,8,30,31], and many articles graphically illustrate this [11,17,37]. Accordingly, we use the area beneath the standard normal distribution to calculate the probability that a selection endpoint lands within a target (Figure 1). If a selection falls beyond $\pm(W / 2)$ from the target center, it falls outside the target and is an error. If it falls within $\pm(W / 2)$, it is a hit. As noted, the speed-accuracy tradeoff assumed by Fitts' law occurs when about 4% of the selection points fall outside $\pm(W / 2)$.

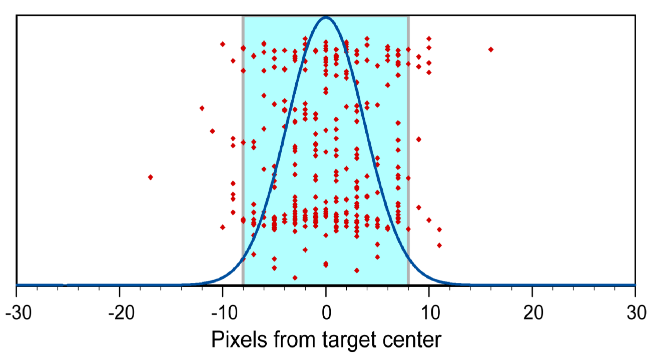

Figure 1. Selection endpoints for a target of size W = 16 pixels. The target is the central highlighted area. Overlaying it is the standard normal distribution, showing ~96% of selections within the target.

The unit-normal deviate, or z-score [17,29], along the standard normal distribution corresponding to a 4% error rate is

$$z = \frac{2.066 \cdot W}{W_e} \tag{6}$$

Note that when $W_e = W$, Eq. 6 equals 2.066, which yields a 4% error rate. Substituting W_e from Eq. 5 yields

$$z = 2.066 \frac{W}{A} \left(2^{\frac{MT_e - a}{b}} - 1 \right) \tag{7}$$

Having obtained the z-score from Fitts' law parameters, we use $\pm z$ to calculate the probability of a selection occurring within that range. The equation for the standard normal distribution $f(x)$ is

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{\frac{-x^2}{2}} \tag{8}$$

The area beneath $f(x)$ from $-z$ to $+z$ gives the probability of a hit within that range. Accordingly, the probability of an error $P(E)$ is 1 minus that quantity

$$P(E) = 1 - \int_{-z}^{+z} f(x)\, dx \tag{9}$$

Expanding $f(x)$ from Eq. 8 and extracting constants gives

$$P(E) = 1 - \sqrt{\frac{2}{\pi}} \int_{0}^{+z} e^{\frac{-x^2}{2}}\, dx \tag{10}$$

Eq. 10 is integrated from 0 to $+z$. This is equal to Eq. 11, which uses the *error function of z*. The error function (*erf*) is a non-elementary function used in probability and partial differential equations. Substituting it yields

$$P(E) = 1 - erf\left(\frac{z}{\sqrt{2}} \right) \tag{11}$$

By substituting Eq. 7 for z in Eq. 11, we arrive at our Fitts-based error model for pointing:

$$P(E) = 1 - erf\left(\frac{2.066 \frac{W}{A} \left(2^{\frac{MT_e - a}{b}} - 1 \right)}{\sqrt{2}} \right) \tag{12}$$

As a check, note that when actual movement time equals the movement time predicted by Fitts' law (i.e., for MT_e in Eq. 12 substitute $a + b \log_2(A/W + 1)$), Eq. 12 reduces to

$$P(E) = 1 - erf\left(\frac{2.066}{\sqrt{2}} \right) \approx 0.0388 \tag{13}$$

Conversely, if $MT_e \neq MT$, the A, W, a, and b terms do not cancel, resulting in error rates other than 4%. This is the issue addressed by Crossman's *post hoc* correction [5]. Further details about the 4% error rate are found elsewhere [17,29,31].

Although there is no closed-form solution for computing the area under the curve in Eq. 10, tables, functions,[1] and approximations [27] are available. With these, we can graph Eq. 12 for different *ID*s using fixed a and b coefficients (Figure 2a), or for different a and b coefficients using a fixed *ID* (Figure 2b).

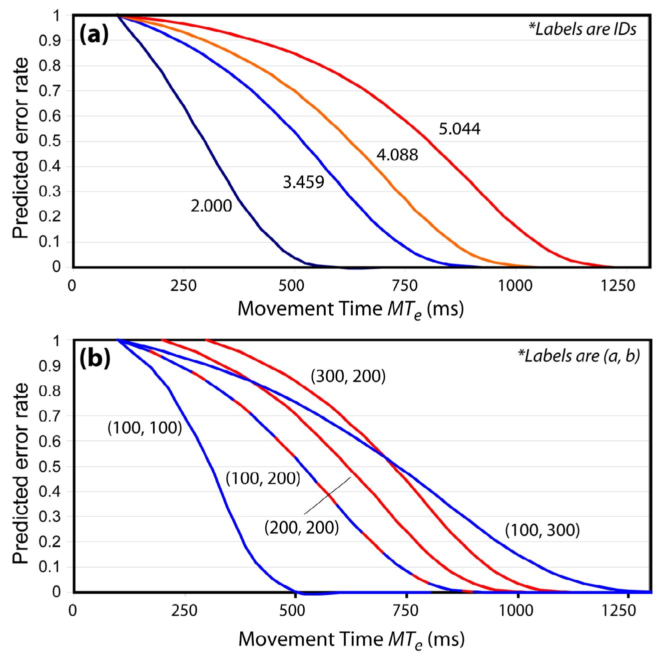

Figure 2. Predicted error rates from Eq. 12 by MT_e: **(a)** for four *ID*s using a = 100 ms, b = 200 ms/bit; and **(b)** for varying combinations of a and b coefficients using *ID* = 3.459 bits. As a increases, the sigmoids translate to the right (+x-axis). As b increases, the sigmoids get "pulled" to the right by their tails.

Derivation Assumptions

As with any model, our Fitts-based error model for pointing (Eq. 12) relies on certain assumptions. One is that Fitts' law holds over a range of movement times (MT_e) even while A and W remain unchanged. For low MT_es where the MT_e / MT ratio[2] is less than, say, 0.60, the kind of rapid aimed movements assumed by Fitts' law where users correct their motion becomes difficult. This is part of the distinction between *closed-loop* and *open-loop* movements [11,17]. The latter are akin to throwing a dart, where an initial ballistic action determines the path of uncorrected motion. In practical terms, this means that participants do not fine-tune the location of the mouse pointer. Instead, they "throw" the cursor at the target. If faster movements are, in fact, open-loop, then we expect the error model to poorly fit our error rates for extreme MT_es. However, prior work [9,37] shows that Fitts' law holds for deliberate or hasty movements. Our experiment indeed confirms this.

---

[1] The ERF function in Microsoft Excel is a useful instance. Some packages lack ERF. For those, note that $1 - \text{ERF}(z/\sqrt{2})$ in Eq. 11 is equivalent to $2(1 - \text{NORMSDIST}(z))$, where NORMSDIST returns the standard normal cumulative distribution function (*cdf*).

[2] This ratio is the observed movement time (MT_e) divided by the movement time predicted by Fitts' law (MT). A value of 0.60 means that when trying to hit a target, participants are taking only 60% of the time that Fitts' law predicts they should take.

Another assumption is that the speed-accuracy tradeoff is logarithmic as Fitts' law purports. There is support for this [6,8,30,31] in spatially constrained time-minimization tasks, like Fitts' reciprocal tapping task. However, other work [26,34,35] shows a *linear* speed-accuracy tradeoff for temporally constrained tasks, like dotting on line targets with a pencil in sync with a metronome. Because our error model (Eq. 12) is derived from Fitts' law, it assumes that closed-loop corrections can be made in accordance with Fitts' logarithmic form [6,22,30]. In our experiment, we are careful to manipulate MT_e so as to maintain Fitts' logarithmic speed-accuracy tradeoff, as we will show.

An additional assumption is that single a and b regression coefficients coincide with changing levels of MT_e, since they remain constant in Eq. 12 while MT_e varies. This is analogous to Fitts' law, where a and b remain constant over a wide range of *ID*s. However, prior work [37] suggests that a and b for different speed-accuracy biases do not fully converge using Crossman's *post hoc* correction [5]. As we demonstrate, single a and b values *do* remain constant in our error model, but they should be elicited for a variety of MT_e values spanning the movement times of interest.

A final assumption is that the selection endpoints are Gaussian about the target center, and that W_e is an accurate reflection of this. Crossman's correction assumes this and prior work confirms this [6,8,30,31]. We further assume that a Gaussian spread occurs when MT_e varies but A and W are fixed. This may be true only for a range of MT_e, since greater kurtosis and skew (the "peakedness" and asymmetry of a distribution) are sometimes observed at particularly fast and slow velocities [12].

Having derived the Fitts-based error model for pointing and made the assumptions explicit, we now review related work addressing the speed-accuracy tradeoff. We then present our experimental findings.

RELATED WORK

Although no prior work has formulated a predictive error rate equation using the parameters of Fitts' law, psychomotor research has been devoted to understanding movement output variability, which is the underlying source of errors. Due to space constraints, we only review seminal examples. For in-depth surveys, readers are directed elsewhere [12,22,25].

The first work often credited with investigating the speed-accuracy tradeoff is Woodworth's in 1899 [33]. By manipulating movement amplitudes and times in a line-drawing task, Woodworth showed that deviations are, in fact, dependent upon movement velocity. But Woodworth never formalized the speed-accuracy tradeoff.

In 1954, Fitts [8] was the first to formalize the speed-accuracy tradeoff in what is now Fitts' law (Eq. 1). Fitts' law does not directly allow for the prediction of error rates, but it mathematically implies the error model we derived (Eq. 12).

In 1963, Crossman and Goodeve [6] proposed the *deterministic iterative-corrections model* to explain Fitts' logarithmic speed-accuracy tradeoff in terms of submovement corrections. Again, no error rate equation was derived, and subsequent work has cast doubt on the underlying claims of this model [22].

In 1979, Schmidt et al. [26] studied how the spread of hits changes with speed in open-loop movements aimed at thin target lines ($W = 0$) and controlled by a metronome. The data of interest were the stylus-mark distributions created around these lines. Schmidt et al. found that the standard deviation of these marks was linearly related to velocity. This led to Schmidt's law:

$$W_e = b\frac{A}{MT_e} \tag{14}$$

Note that for Schmidt et al., the standard deviation of hits W_e was *not* Crossman's corrected W_e, which is actually 4.133 times this value. Regardless, note how Schmidt's linear relationship differs from Eq. 5 where the relationship is logarithmic. We can arrange Schmidt's law with MT_e on the left, contrasting this to Fitts' law (Eq. 2):

$$MT_e = b\frac{A}{W_e} \tag{15}$$

As a result of Schmidt's work, researchers learned that spatially-constrained ($W > 0$) time-minimization tasks follow Fitts' law, while temporally-constrained tasks without spatial constraints ($W = 0$) follow Schmidt's law [20,34]. To examine the confluence of these issues, Zelaznik et al. [35] constrained both time *and* space, finding that strict temporal constraints result in a linear speed-accuracy tradeoff even when $W > 0$. However, like Schmidt et al., Zelaznik et al. used a stylus, which inextricably bounds *arrival* at an endpoint with *selection* at that endpoint. In our experiment, although we also used a metronome, we retained a logarithmic speed-accuracy tradeoff. This is likely because participants were explicitly instructed to *click* with the metronome but not to *move* with it. This allowed them to arrive at the target and correct the mouse position before selection, if necessary. We return to this issue in our experimental results.

In 1988, Meyer et al. [21] sought to accommodate Fitts' logarithmic model and Schmidt's linear model with the *optimized dual-submovement model*, which held that the error in a submovement is proportional to its velocity. This model allowed for two submovements: an initial ballistic one and an optional corrective one. Movement time could be predicted with Eq. 16:

$$MT = a + b \cdot \sqrt[2]{\frac{A}{W}} \tag{16}$$

Unlike most prior work, Meyer et al. actually formulated a predictive error equation:[3]

$$P(E) = 2 \cdot (1-c) \cdot \left(1 - N \Big/ \sqrt{\theta \frac{A}{W} - 1}\right) \tag{17}$$

For our purposes, the important point in Eq. 17 is that it lacks an explicit term for time, making it suitable only for tasks where participants move at the model's predicted *MT*, making it of limited value for our current investigation.

Interestingly, Meyer et al. [22] (pp. 213-215) discovered later that when an infinite number of submovements are permitted, their dual model (Eq. 16) results in a very familiar functional form!

$$\lim_{n\to\infty} a + b \cdot \sqrt[n]{\frac{A}{W}} = a' + b' \cdot \log_e\left(\frac{A}{W}\right) \tag{18}$$

In 1997, Plamondon's *kinematic theory* [25] used a delta-lognormal law to generate velocity profiles for both open- and closed-loop movements. Although this approach fit a variety of velocity profiles, it was criticized for lacking an explanation of underlying dynamics. As with most prior work, no predictive error equation was derived.

In 2004, Zhai et al. [37] showed that Crossman's *post hoc* correction [5] does not fully compensate for differences in participants' speed-accuracy biases. It does, however, help *a* and *b* to converge. Zhai et al. showed that when using the correction, *a* and *b* should be elicited over all speed-accuracy levels, not just a subset, a finding we reinforce here. Zhai et al. also found disproportionate effects of *W* concerning the spread of hits, another finding we replicate. No error rate equation was proposed.

In 2005, Grossman and Balakrishnan [11] extended Fitts' law to two dimensions by numerically mapping the probability of an open-loop movement hitting a 1D target to Fitts' *ID*. Using this mapping, they generated *ID*s for 2D targets based on the chance that an open-loop movement lands inside a 2D region. Although they used the concept of the probability of a hit, they did not define an error equation or use MT_e as a parameter affecting this probability.

Clearly, a great deal of work on the speed-accuracy tradeoff for aimed movements exists, but few efforts have formulated predictive error rate equations. We now describe an experiment to test the validity of our error model for pointing.

EXPERIMENT

The experiment controlled target size, target distance, and movement time in a one-dimensional reciprocal pointing task. The goal was to examine the correctness of Eq. 12 for predicting errors based on Fitts' law parameters.

Method

Participants

Sixteen participants, ages 19-41, volunteered for the study. All were right-handed computer users. They were compensated with a voucher for food at a local cafeteria.

Apparatus

Testing was conducted with a 17" LCD monitor set to 1280×1024 resolution and connected to a Compaq EVO desktop computer running Windows XP (Pentium 4, 2.2 GHz processor, 1 GB RAM). The same IntelliMouse Optical was used by all participants, with its speed set to 7/10 in the mouse control panel and acceleration turned off. Software was authored in C# using .NET 2.0 and presented trials to participants while logging their mouse activities in XML. The software ran full-screen. All other applications and nonessential services were disabled.

Procedure

The study had two phases. In the first phase, participants performed a conventional Fitts' reciprocal pointing task to elicit their personal Fitts' law models. Each participant performed 12 target acquisitions for 3 target sizes (*W*: 16, 32, 64 px) × 3 target distances (*A*: 192, 320, 512 px), which comprised 9 distinct *ID*s ranging from 2.000 to 5.044 bits. Participants repeated these conditions twice for a total of 12 × 9 × 2 = 216 target acquisitions. With 16 participants, this resulted in 3456 total acquisitions.

If a target was missed, it flashed red and an error sound was played. The error rate, which participants were to maintain at 4%, was shown after each *A*-*W* condition. At the end of the first phase, the software performed a Fitts' law analysis, displaying *a* and *b* coefficients for that participant. These coefficients were used in the second phase.

In the second phase, a visual and auditory metronome was used to manipulate participants' movement times (MT_e). We did not control movement time explicitly using raw time values (e.g., 100 ms, 200 ms, 300 ms) because doing so ignores individual participant differences—what is "fast" to one participant may be "slow" to another. Instead, we used each participant's *a* and *b* coefficients from the first phase to determine the *MT* predicted by Fitts' law, and then we set the nominal metronome time, MT_m, based on a percentage of *MT*, which we call $MT_\%$. So the metronome time was

$$MT_m = MT_\% \times MT \tag{19}$$

and the manipulated independent variable, $MT_\%$, was

$$MT_\% = \frac{MT_m}{MT} \tag{20}$$

Thus, when $MT_\% < 1.00$, participants moved faster than Fitts' law predicts. When $MT_\% > 1.00$, they moved slower than Fitts' law predicts. But the actual MT_m values were tailored to each participant.

[3] In Eq. 17, *c* is a constant and $N/\sqrt{[\theta\cdot(A/W) - 1]}$ is a probability related to the standard normal distribution. Interested readers are referred to pp. 349-350 for the mathematical details [21].

Ten levels of $MT_{\%}$ were used, from 0.40 to 1.30 stepping every 0.10. The values for A and W were the same as in phase one. Thus, there were 90 $A \times W \times MT_{\%}$ combinations. For each $A \times W \times MT_{\%}$ condition, participants performed 12 target acquisitions, with the first two ignored as practice. The error rate for the 10 remaining target acquisitions was measured as the outcome for each condition. Participants were tested over the 90 conditions twice, resulting in 180 error rate measures from 1800 clicks per participant. With 16 participants, the experiment consisted of 2880 error rate measures from 28,800 clicks.

The metronome had both visual and auditory components. The same vertical targets in phase one were used in phase two, but in phase two they were outlined by a gray animated border that grew smoothly from the top and bottom toward the center (Figure 3). When the borders met in the middle, a "tick" sound was played, and the borders disappeared, only to begin growing again from the target ends. Participants found this feedback to be clear, as it allowed them to both see and hear the progress of the metronome.

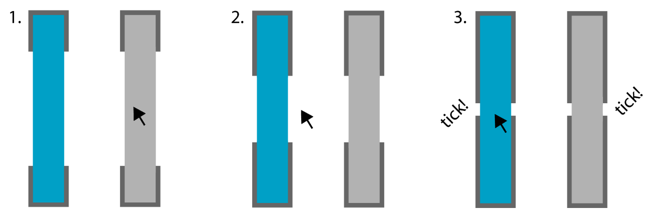

Figure 3. Vertical bar targets in the metronome phase of the study had animated borders that grew towards the center. (Targets, distances, and the mouse cursor are not shown to scale.)

Participants were instructed to be as accurate as possible while clicking in sync with the metronome. They were not told to *move* with the metronome, but, rather, to *click* with it. Although prior studies [26,34,35] of temporally constrained movements show a linear speed-accuracy tradeoff, participants moved in tandem with a metronome, usually dotting between lines with a stylus. As our results show, the data indeed exhibit the logarithmic speed-accuracy tradeoff predicted by Fitts' law (Eq. 1) and not the linear tradeoff shown by Schmidt et al. (Eq. 14).

Results

Adjustment and Analysis of Data

In both phases of the experiment, Crossman's correction [5] was used. This usually lowers R^2 but provides more appropriate a and b coefficients [37]. Consistent with prior work [16], spatial outliers in phase one were removed if the observed movement amplitude was less than half the nominal movement amplitude, or if the selection endpoint was more than twice the target width from the target center. This removed only one trial from phase one (0.03%). In phase two, spatial outliers were *not* removed because they could constitute legitimate errors. Instead, temporal outliers were removed if participants deviated considerably from the prescribed movement time MT_m. A temporal outlier occurred when an observed movement time MT_e was less than 75% or more than 150% of the metronome time MT_m. This removed 308 trials (1.07%) from phase two. In both phases, Fitts' law models were fit for each participant using means from each $A \times W$ or $A \times W \times MT_{\%}$ condition, with repeated conditions remaining separate ($N = 18$ points or $N = 180$ points per participant in each phase, respectively). For statistical tests, our error rate data are not normally distributed and do not allow ANOVA tests. Accordingly, we use nonparametric Wilcoxon sign-rank and rank-sum tests [32].

Calibration Phase: A × W

For phase one, average MT_e was 710 ms with an average error rate of 4.37%. The mean TP was 5.08 bits/s, which is close to results from prior studies [16]. The average a and b coefficients were 94.21 ms and 173.64 ms/bit, respectively. The average per-participant fit of Fitts' law was $R^2 = .763$ for $N = 18$ points. Although this is lower than often reported for Fitts' law studies, most studies fit fewer data points than we do here. The goal of phase one was only to elicit Fitts' a and b parameters for each participant so that $MT_{\%}$ could be used in phase two.

Metronome Phase: A × W × MT%

For phase two, movement times (MT_e) ideally should match the nominal metronome times (MT_m). Average MT_e was 613 ms, and the average MT_e / MT_m ratio was 1.04, indicating compliance with the metronome. The average error rate was 21.24%. The average throughput was 5.42 bits/s. The average a and b coefficients over all levels of $MT_{\%}$ were -84.22 ms and 224.61 ms/bit, respectively. The average per-participant fit of Fitts' law was $R^2 = .766$ for $N = 180$ points.

Having controlled MT_e using a metronome, it is important to verify that our trials exhibited Fitts' logarithmic speed-accuracy tradeoff and not the linear tradeoff of Schmidt et al. Figure 4a shows that Fitts' law fits our data very well ($R^2 = .999$). If a linear tradeoff were in effect, Figure 4b would show a straight line, not a curve. In Figure 4c, we see that Schmidt's law (Eq. 14) does not hold very well ($R^2 = .601$). In addition, we performed these analyses just on fast trials with $MT_e < 500$ ms and found the same results. Thus, our experimental manipulation of MT_e maintains the logarithmic speed-accuracy tradeoff as modeled by Fitts' law.

Another check is whether the observed error rate for $MT_{\%} = 1.00$ was near the 4% nominal error rate assumed by Fitts' law. In fact, the error rate for $MT_{\%} = 1.00$ was 5.78%, which is reasonably close to the ideal value of 4%. Thus, despite the manipulation of MT_e with the metronome, participants performed according to Fitts' law.

Overall Model Fit

Having confirmed that the experimental manipulations were sound, we turn to the performance of our error model for pointing (Eq. 12). Figure 5a shows the $N = 90$ points for each combination of $A \times W \times MT_{\%}$ plotted as (*predicted*, *observed*) ordered pairs. A "perfect model"

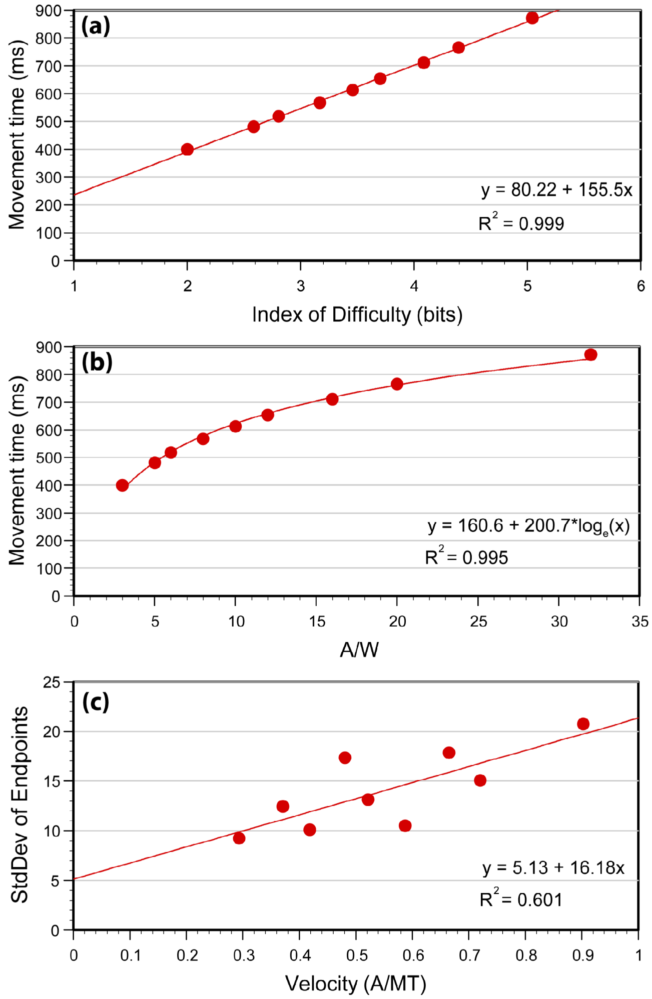

Figure 4. (a) Fitts' law holds well for the metronome trials. **(b)** If Schmidt's law (Eq. 15) were fitting, these points would fall in a line. **(c)** Schmidt's law (Eq. 14) does not offer a strong fit for this data.

would place every point on the diagonal ($y = x$). Linear regression results in $y = 0.007 + 0.958\,x$ ($R^2 = .959$), which has a near-zero intercept and near-unity slope. The model fit is significant ($F_{1,88} = 2081.14$, $p < .0001$). On average, predicted error rates are within 3.59 percentage points of observed error rates and are not significantly different ($z = 74.50$, $p = .77$). Figure 5b shows results for each level of $MT_{\%}$ ($R^2 = .992$).

Effect of MT%

Our manipulation of $MT_{\%}$ allows us to examine how error rates vary with movement time. Not surprisingly, there was a significant effect of $MT_{\%}$ on error rate ($\chi^2_{(9,N=90)} = 74.40$, $p < .0001$). Figure 6a shows predicted and observed error rates for each $A \times W \times MT_{\%}$ point over $MT_{\%}$. Predicted and observed error rates show strong correlations with $MT_{\%}$ ($R^2 = .767$ and $R^2 = .717$, respectively). Figure 6b shows the same data averaged over *ID* for each level of $MT_{\%}$. Again, predicted and observed rates strongly correlate with $MT_{\%}$ ($R^2 = .923$ and $R^2 = .883$, respectively).

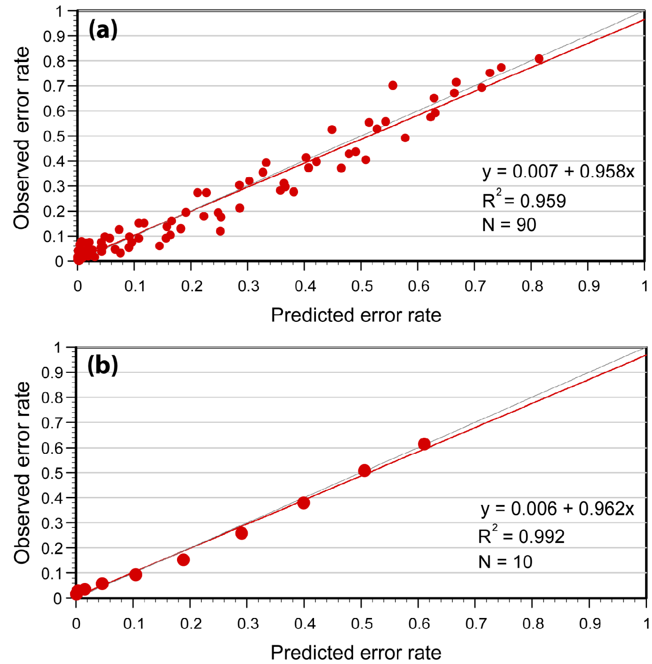

Figure 5. (a) Observed error rates *vs.* predicted error rates for each *A*×*W*×*MT%*. **(b)** Averaged over *ID* for each level of $MT_{\%}$.

Effect of A, W, and ID

Owing to its origins in Fitts' law, the error model for pointing maintains that target distance (*A*) and size (*W*) contribute proportionally to predicted error rates. However, as we now discuss, this is not the case with our data.

Average error rates for increasing levels of *A* were 19.59%, 21.13%, and 23.62%, respectively. These differences did not constitute a significant effect on error rate ($\chi^2_{(2,N=90)} = 0.15$, n.s.). Figure 7a recasts Figure 6a with data points grouped by *A*. An intermixing of error rates is evident, consistent with Fitts' notion of *ID*.

However, the same is *not* true of *W*. Average error rates for decreasing levels of *W* were 12.37%, 20.97%, and 30.99%, respectively. Unlike levels of *A*, these differences exert a significant effect on error rates ($\chi^2_{(2,N=90)} = 11.73$, $p < .01$). Figure 7b recasts Figure 6a with data points grouped by *W*. Bands are clearly visible for each level of *W* without the intermixing consistent with Fitts' notion of *ID*.

In Figure 8a, we group the data in Figure 4a by *W*. Movement time is affected by *W* just as Fitts' law predicts, namely in combination with *A* as *ID*. Figure 8b shows how this corresponds to predicted error rates: As expected, the graphs look almost identical. However, in Figure 8c, when we plot the actual observed error rates, a discontinuity appears: Decreasing *ID*, which should cause error rates to go down, actually causes error rates to go up when target size (*W*) decreases. These findings indicate that the unified notion of *ID* does not hold for errors as it does for movement time. This incongruence seems systematic, however, as Figure 8d shows that error rate over-estimation is higher for larger *W*, and lower for smaller *W*. Further research is necessary to refine the role of *W* for affecting

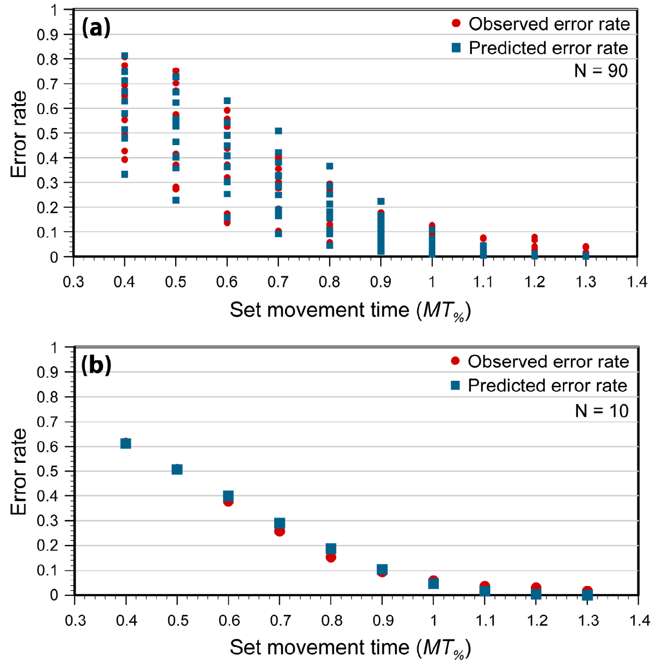

Figure 6. (a) Predicted and observed error rates for each $A \times W \times MT\%$ across levels of $MT_{\%}$. **(b)** Averaged over *ID* for each level of $MT_{\%}$.

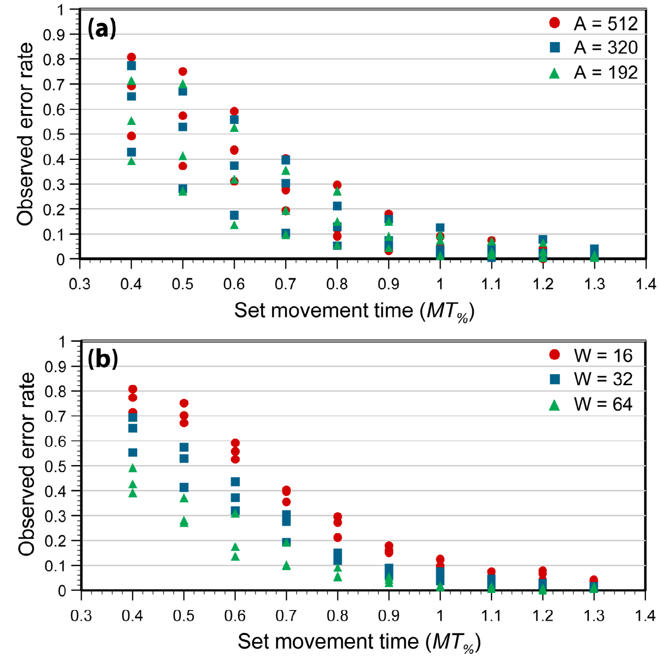

Figure 7. (a) Observed error rates grouped by levels of *A* show an intermixing of rates consistent with *ID* and the error model. **(b)** When grouped by levels of *W*, banding appears, indicating error rates are overly influenced by *W* compared to *A*.

pointing error rates; from these results, it is clear that *W* and *A* do not contribute proportionally, indicating a departure from Fitts' law.

DISCUSSION

On the whole, our error model for pointing provides good error-rate predictions (Figure 5). Although models are always imperfect and measurements noisy, the match between error rate predictions and observed error rates is strong, especially given the contrived nature of experimentally controlled movement times. Our data confirm the logarithmic speed-accuracy tradeoff and the relative harmony between our metronome-guided pointing and pointing modeled by Fitts' law (Figure 4). This itself is noteworthy, as prior metronome studies [26,34,35] often exhibit a linear speed-accuracy tradeoff.

As Figure 2b shows, the values for Fitts' *a* and *b* coefficients substantially affect predicted error rates. Future work is necessary to tease out the sensitivities of the error model to its parameters. Our own explorations indicate that *a* and *b* should be elicited from trials that span the movement times of interest. We discovered that *single* per-participant values for *a* and *b* improve model predictions compared to separate *a* and *b* values for each level of $MT_{\%}$, or values from a subset of $MT_{\%}$s. These insights are consistent with Zhai et al. [37]. Also, *a* and *b* values from traditional Fitts tasks, where error rates are held at 4% (e.g., phase one), make for poor error model predictions. Instead, when eliciting *a* and *b* for use with the error model, Crossman's correction should be applied, all movement time conditions should be pooled, and per-participant *a* and *b* should be elicited.

It is important to emphasize that neither the metronome nor the notion of "$MT_{\%}$" is necessary for *a* and *b* elicitation; these were only used for experimental manipulation. Likewise, for fitting an instance of the error model, it is not necessary to run the traditional Fitts calibration trials from phase one. The error model for pointing (Eq. 12) uses raw time values (MT_e), so it is only necessary to use trials that cover a range of MT_e along with varied *A* and *W*. In other words, a researcher wishing to successfully employ the error model only needs to manipulate *A* and *W* (as with any Fitts study) and elicit movements that cover a range of slow and fast speeds (MT_e).

W's disproportionate effect on error rates is important because it is at odds with Fitts' notion of *ID*. Zhai et al. [37] had similar findings: Large target widths are under-utilized, while small target widths are over-utilized. Even *within* the same level of *ID*, Wallace and Newell [30] found lower error rates for larger *W*. And C. L. MacKenzie et al. [14] found that velocity profiles are determined much more by *W* than by *A*, which may be a clue. Clearly, more work is necessary to refine *W*'s role in the error model for pointing.

Finally, we should note that an error model is useful in areas of human factors outside HCI. For example, on an assembly line, inspectors might have limited time to grab items as they pass by. And the design of aircraft cockpit controls, with which Fitts himself was quite familiar [7], might be informed by better error prediction and estimation.

FUTURE WORK

As noted, further work should tease out the model's sensitivity to the *a* and *b* coefficients, discovering exactly

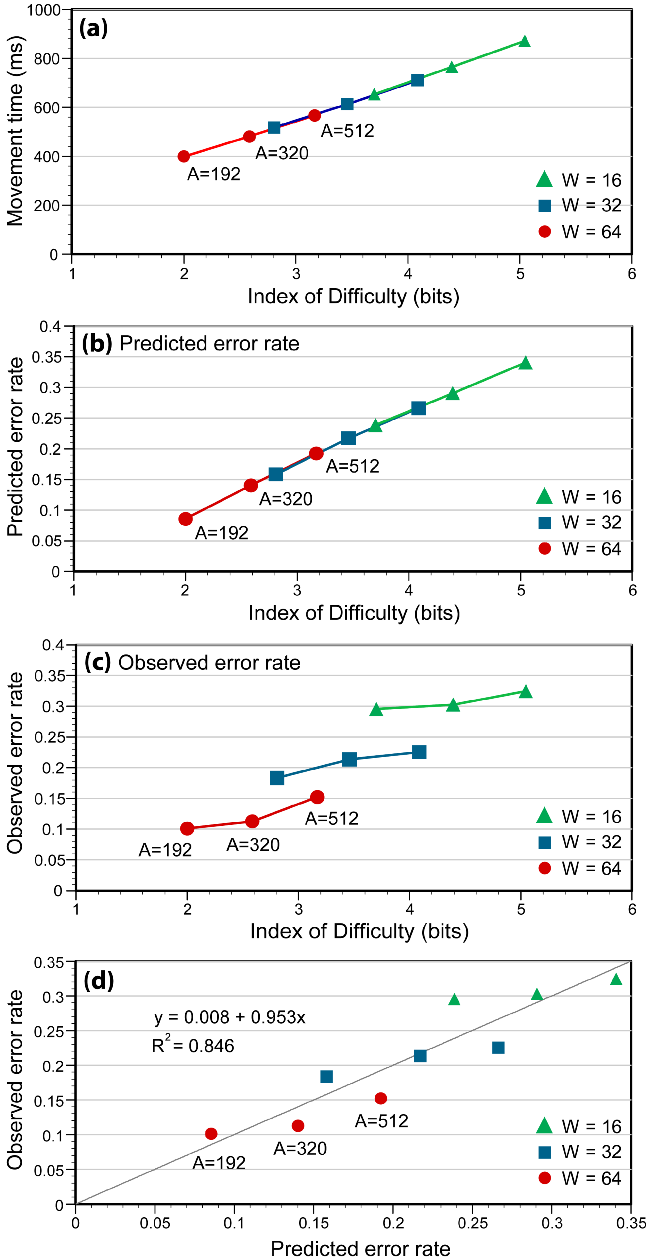

Figure 8. (a) Movement time follows Fitts' law according to *ID*, with proportional influence from *W* and *A*. **(b)** Error model predictions follow suit, owing to their basis in Fitts' law. **(c)** Observed error rates indicate a discontinuity, where lower *ID*s with smaller targets have higher error rates than the model predicts. **(d)** Error rate over-estimation is higher for larger *W* and lower for smaller *W*.

how their elicitation affects model performance. Also, we should discover more precisely the role of *W* and its relationship to *A* in determining pointing errors; clearly the idea of "equal but inverse contribution," so firmly rooted in Fitts' law, does not entirely apply. Other future work should test the model in different experimental conditions, where MT_e is not controlled by a metronome, but instead, perhaps, by different payment schemes [9] or reinforcement [37]. The model also should be tested for discrete movements, rather than Fitts' reciprocal tapping. Also, the model should be tested with a stylus, where *arrival* at a target and *selection* of that target are bound. If a metronome is used with a stylus, it may be difficult to maintain Fitts' logarithmic speed-accuracy tradeoff; perhaps a Schmidt-based error model for pointing could be similarly effective. (The same mathematical process applied to Eq. 5 could be applied to Eq. 14.)

CONCLUSION

The field of human-computer interaction has benefited over the years from quantitative models of human performance, and Fitts' law is undoubtedly the most prevalent of these. However, Fitts' law is centrally concerned with movement-time prediction, not the prediction of error rates. In this work, we "round out" the theory by deriving an error model for pointing that is strongly implied by Fitts' law. The model holds over a range of target sizes, target distances, and movement times, although discontinuities with Fitts' law emerge concerning the role of target size. Researchers, modelers, designers, and usability experts may benefit from quantitative models such as ours, which provide input for design and support rigorous evaluation of interactive systems.

ACKNOWLEDGEMENTS

We thank Patrick Baudisch, Paige Bills, Krzysztof Z. Gajos, Yves Guiard, Ken Hinckley, Scott E. Hudson, Daniel S. Weld, and Shumin Zhai for helpful conversations. This work is supported in part by Intel Research.

REFERENCES

1. Balakrishnan, R. and MacKenzie, I.S. (1997) Performance differences in the fingers, wrist, and forearm in computer input control. *Proc. CHI '97*. New York: ACM Press, 303-310.
2. Balakrishnan, R. (2004) "Beating" Fitts' law: Virtual enhancements for pointing facilitation. *Int'l J. of Human-Computer Studies 61* (6), 857-874.
3. 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.
4. Card, S.K., Moran, T.P. and Newell, A. (1983) *The Psychology of Human-Computer Interaction*. Hillsdale, NJ: Lawrence Erlbaum.
5. 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.
6. 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.
7. Fitts, P.M., Jones, R.E. and Milton, J.L. (1950) Eye movements of aircraft pilots during instrument-landing approaches. *Aeronautical Engineering Review 9* (2), 24-29.

8. Fitts, P.M. (1954) The information capacity of the human motor system in controlling the amplitude of movement. *J. of Exp. Psychology 47* (6), 381-391.
9. 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.
10. Grossman, T. and Balakrishnan, R. (2005) The Bubble Cursor: Enhancing target acquisition by dynamic resizing of the cursor's activation area. *Proc. CHI '05*. New York: ACM Press, 281-290.
11. Grossman, T. and Balakrishnan, R. (2005) A probabilistic approach to modeling two-dimensional pointing. *ACM Transactions on Computer-Human Interaction 12* (3), 435-459.
12. Hancock, P.A. and Newell, K.M. (1985) The movement speed-accuracy relationship in space-time. In *Motor Behavior: Programming, Control, and Acquisition*, H. Heuer, U. Kleinbeck, and K.-H. Schmidt (eds). New York: Springer-Verlag, 153-188.
13. John, B.E. and Kieras, D.E. (1996) The GOMS family of user interface analysis techniques: Comparison and contrast. *ACM Transactions on Computer-Human Interaction 3* (4), 320-351.
14. MacKenzie, C.L., Marteniuk, R.G., Dugas, C., Liske, D. and Eickmeier, B. (1987) Three-dimensional movement trajectories in Fitts' task: Implications for control. *Quarterly J. of Exp. Psychology 39A* (4), 629-647.
15. MacKenzie, I.S. (1989) A note on the information-theoretic basis for Fitts' law. *J. of Motor Behavior 21*, 323-330.
16. MacKenzie, I.S., Sellen, A. and Buxton, W. (1991) A comparison of input devices in elemental pointing and dragging tasks. *Proc. CHI '91*. New York: ACM Press, 161-166.
17. MacKenzie, I.S. (1992) Fitts' law as a research and design tool in human-computer interaction. *Human Computer Interaction 7* (1), 91-139.
18. MacKenzie, I.S. and Buxton, W. (1994) The prediction of pointing and dragging times in graphical user interfaces. *Interacting with Computers 6* (2), 213-227.
19. MacKenzie, I.S. and Soukoreff, R.W. (2003) Card, English, and Burr (1978)—25 years later. *Extended Abstracts CHI '03*. New York: ACM Press, 760-761.
20. Meyer, D.E., Smith, J.E.K. and Wright, C.E. (1982) Models for the speed and accuracy of aimed movements. *Psychological Review 89* (5), 449-482.
21. Meyer, D.E., Abrams, R.A., Kornblum, S., Wright, C.E. and Smith, J.E.K. (1988) Optimality in human motor performance: Ideal control of rapid aimed movements. *Psychological Review 95* (3), 340-370.
22. Meyer, D.E., Smith, J.E.K., Kornblum, S., Abrams, R.A. and Wright, C.E. (1990) Speed-accuracy tradeoffs in aimed movements: Toward a theory of rapid voluntary action. In *Attention and Performance XIII*, M. Jeannerod (ed). Hillsdale, NJ: Lawrence Erlbaum, 173-226.
23. Newell, A. and Card, S.K. (1985) The prospects for psychological science in human-computer interaction. *Human Computer Interaction 1* (3), 209-242.
24. Pavlovych, A. and Stuerzlinger, W. (2004) Model for non-expert text entry speed on 12-button phone keypads. *Proc. CHI '04*. New York: ACM Press, 351-358.
25. Plamondon, R. and Alimi, A.M. (1997) Speed/accuracy trade-offs in target-directed movements. *Behavioral and Brain Sciences 20* (2), 279-349.
26. Schmidt, R.A., Zelaznik, H., Hawkins, B., Frank, J.S. and Quinn, J.T. (1979) Motor-output variability: A theory for the accuracy of rapid motor acts. *Psychological Review 86* (5), 415-451.
27. Shah, A.K. (1985) A simpler approximation for areas under the standard normal curve. *The American Statistician 39* (1), 80.
28. Soukoreff, R.W. and MacKenzie, I.S. (1995) Theoretical upper and lower bounds on typing speed using a stylus and soft keyboard. *Behaviour and Information Technology 14* (6), 370-379.
29. Soukoreff, R.W. and MacKenzie, I.S. (2004) Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research in HCI. *Int'l J. of Human-Computer Studies 61* (6), 751-789.
30. Wallace, S.A. and Newell, K.M. (1983) Visual control of discrete aiming movements. *Quarterly J. of Exp. Psychology 35A* (2), 311-321.
31. Welford, A.T. (1968) *Fundamentals of Skill*. London: Methuen.
32. Wilcoxon, F. (1945) Individual comparisons by ranking methods. *Biometrics Bulletin 1* (6), 80-83.
33. Woodworth, R.S. (1899) The accuracy of voluntary movement. *Psychological Review, Monograph Supplements 3* (2, Whole No. 13), 1-114.
34. Wright, C.E. and Meyer, D.E. (1983) Conditions for a linear speed-accuracy trade-off in aimed movements. *Quarterly J. of Exp. Psychology 35A* (2), 279-296.
35. Zelaznik, H.N., Mone, S., McCabe, G.P. and Thaman, C. (1988) Role of temporal and spatial precision in determining the nature of the speed-accuracy trade-off in aimed-hand movements. *J. of Exp. Psychology: Human Perception and Performance 14* (2), 221-230.
36. Zhai, S., Hunter, M. and Smith, B.A. (2000) The Metropolis keyboard—An exploration of quantitative techniques for virtual keyboard design. *Proc. UIST '00*. New York: ACM Press, 119-128.
37. 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'l J. of Human-Computer Studies 61* (6), 823-856.