

The Effect of Time-based Cost of Error in Target-directed Pointing Tasks

Nikola Banovic^{1,2}, Tovi Grossman¹, and George Fitzmaurice¹

¹Autodesk Research

210 King St. East, Toronto, Ontario, Canada, M5A 1J7
{firstname.lastname}@autodesk.com

²HCI Institute, Carnegie Mellon University

5000 Forbes Ave., Pittsburgh, PA 15213
nbanovic@cs.cmu.edu

ABSTRACT

One of the fundamental operations in today's user interfaces is pointing to targets, such as menus, buttons, and text. Making an error when selecting those targets in real-life user interfaces often results in some cost to the user. However, the existing target-directed pointing models do not consider the cost of error when predicting task completion time. In this paper, we present a model based on expected value theory that predicts the impact of the error cost on the user's completion time for target-directed pointing tasks. We then present a target-directed pointing user study, which results show that time-based costs of error significantly impact the user's performance. Our results also show that users perform according to an expected completion time utility function and that optimal performance computed using our model gives good prediction of the observed task completion times.

Author Keywords

Fitts' law; movement time; pointing time; pointing errors; error cost; speed-accuracy tradeoff.

ACM Classification Keywords

H.5.2. [Information interfaces and presentation]: User Interfaces – Theory and methods, evaluation/methodology.

INTRODUCTION

One of the fundamental operations in today's user interfaces is pointing to targets, such as menus, buttons, and text. Fitts' law [9, 20] is arguably one of the most influential quantitative methods that predict the time required to acquire such targets. Traditionally, Fitts' law experiments studied target-directed pointing tasks without considering the effect of the cost of an error when the user misses the target. In real-life, however, when the user makes a selection error there is often some cost associated with that error. The user needs to recover from the error, which requires time, and then continue until she selects the target. For example, if the user clicks on the wrong menu item in the application menu, the user has to undo the effects of the wrong command, but also has to navigate the menu again in order to select the correct menu item.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

CHI 2013, April 27–May 2, 2013, Paris, France.

Copyright © 2013 ACM 978-1-4503-1899-0/13/04...\$15.00.

As the cost of error increases, psychology research [25] may indicate that the user will change their behavior in order to reduce the error rate. The speed-accuracy tradeoff [29] suggests that in order to reduce the error rate, the user will have to move slower which will increase their task completion time. Users will likely favor efficiency [12, 13, 30] and tend towards optimal speed-accuracy tradeoff which will minimize their task completion time. However, although there is extensive research in speed-accuracy tradeoff (see [15, 24, 26] for reviews), there is little research that explores how the cost of error impacts the user's behavior and performance in target-directed pointing tasks. It is not clear how much slower or how much more accurately the user will perform the task as the cost of error increases, and whether the user will converge towards an optimal behavior.

In this paper, we present a model that predicts the impact that the cost of making an error will have on the user's target-directed pointing task completion time. We base our model on the expected utility theory [28], where we assume that the user will favor those strategies which minimize the expected task completion time. Similar to previous studies [25], in our model we abstract the cost of error as a time-based penalty, which represents the time required to recover from the error. Our model is not intended to replace Fitts' law. Instead, it builds on top of Fitts' law [9, 20] in conjunction with Wobbrock *et al.*'s error model [34, 35], to calculate an optimal movement time the user should attempt to acquire a target at, taking into account the time-based penalty associated should the user miss.

After describing our model, we present a target-directed pointing user study, which shows that time-based penalties significantly impact the users' completion time (the time until a successful target selection), movement time (the time until the first target selection attempt), and error rate. Furthermore, the results show that optimal performance computed using our model predicts the observed task completion and movement time well, but has room for improvement when predicting the error rate.

The results of our study indicate that our model can be used to predict the target-directed pointing task completion times in user interfaces for different time-based costs of error. This will allow both researchers and designers to model the user performance with a given interface in practice. We conclude with discussion of the implications of our findings on different parameters of the target-directed pointing task.

RELATED WORK

Cognitive modeling methods [4] are often used to predict the time it takes to perform different tasks when interacting with user interfaces. One way to predict the user's target-directed pointing time in the user interface in those cognitive modeling methods is using the Fitts' law [9, 20]. Although, Woodworth [37] is often credited as the first to propose a relationship between movement time, amplitude, and velocity in determining movement accuracy in target-directed pointing tasks, it was Fitts [9] who first formalized the speed-accuracy tradeoff. For the sake of brevity, we refer the reader to in-depth surveys for detailed review of research on speed-accuracy tradeoff [15, 24, 26].

The cognitive modeling methods mentioned above primarily focus on error-free performance. As such, those methods do not predict the time it takes to correct an error; instead they use observed error recovery times for the particular task. The error rates used to calculate the expected completion times in those models are often calculated as probability of unconscious or automatic errors [27], and not errors caused by the speed-accuracy tradeoff.

Existing target-directed pointing studies often do not consider the cost of error when predicting task completion time because those experiments focus on comparing throughput of different pointing devices (e.g., [6]), and use only error-free performance with the devices. However, a task completion time model that considers the cost of error could better explain the impact that the cost could have on the user's behavior in practice.

Given a risk of making an error, users are likely to favor target-directed pointing microstrategies that allow for higher levels of accuracy [25]. Furthermore, Soukoreff and MacKenzie [30] argued that the users have freedom of choice over the characteristics of their movement, and also have a preference for efficiency. Elliot *et al.* [8] showed that users are able to learn to optimize their performance over the course of performing a target-directed pointing task. In fact, past research [12, 13] showed that users change their target-directed pointing behavior in response to even small time savings at the level of milliseconds. Furthermore, Gray and Boehm-Davis [12] proposed that the choice of these microstrategies is non-deliberate. These previous findings have not considered the cost of making an error during the pointing task, but do serve as motivation for our work.

Although some studies have explored target-directed movement at pace naturally selected by the user (e.g., [1, 38]), most studies motivate their users in order to study the users' maximum capacities. In studying speed-accuracy tradeoffs, past research primarily used explicit verbal instructions of how fast or how accurately participants should move (e.g., maintain 4% error rate [29]), often including conditions at different levels of speed-accuracy tradeoff (e.g., performing as fast as possible disregarding any errors or performing with no errors [14, 21, 40]). Those

works confirm the user's ability to modify characteristics of their movement to accommodate different levels of speed-accuracy tradeoff when instructed to do so.

However, users rarely receive explicit instructions on how to manipulate their speed-accuracy tradeoff in real-life. Often other factors such as incentives to complete the task faster, or penalties related to incorrectly selecting the target implicitly affect the user's speed-accuracy tradeoff. Fitts and Radford [10] and Elliott *et al.* [8] showed that users adjust properties of their movement based on different monetary compensation levels associated with faster or more accurate movement. Similarly, Al-Imam and Lank [2] showed that monetary incentive-based, as well as penalty-based approaches, could affect the user's speed-accuracy tradeoff. We build on top of such research, but we specifically focus on how time-based penalties for inaccurate movements affect the user's speed-accuracy tradeoff. We also extend those works by presenting a model that predicts the impact of time-based penalties on the user's task completion time.

Recent models for optimal movement planning (e.g., [17, 31, 32]) predict optimal movement endpoints based on the constraints of the task. However, our model does not provide an *a priori* prediction of the completion time in target-directed pointing tasks. Instead, our work extends the Fitts' law, as traditionally used in HCI research for empirical estimate of task completion times, to include the cost of error. This is similar to models by Dean *et al.* [6] and Hudson *et al.* [18]. However, instead of imposing penalties on early or late selection times [18] or using explicit decay of an incentive to represent the cost of moving slow to select the target [6], we abstract the cost of error as a fixed amount of time that only occurs when the user misses the target. Thus, we allow the user to decide their tradeoff between the timesaving when moving fast to select the target and the time it would take to recover from an error should the user miss the target.

Post hoc corrections [5, 20, 29, 33] could be another way to relate user performance at different time-based penalty levels. However, those corrections exist specifically to normalize the data between speed-minded and accuracy-minded participants. Although they allow for data aggregation, they do not explain why this difference exists, or why some users choose different strategies. More importantly, they do not suggest how time penalty associated with making an error affects the motivated user.

Error models based on Fitts' law [23, 34, 35] allow for predicting error rates based on parameters of a Fitts' law task. Meyer *et al.* [23] formulated an error model which relates the user's error rate and the model's predicted movement time. Similarly, Wobbrock *et al.* proposed a 1-dimensional [34] and 2-dimensional [35] target-directed pointing error models that relate the user's effective movement time with the user's error rate. The explicit term for movement time in those two models [34, 35] makes

them suitable for estimating error rates when the user moves at movement times different from the Fitts' law model's predicted movement time. Therefore, we use Wobbrock *et al.*'s 2-dimensional error model [35] in our cost-based task completion time model.

COST-BASED COMPLETION TIME MODEL

Making an error when selecting targets, such as menus, buttons, and text in actual user interfaces often results in some cost to the user. In order to predict performance of an interface, a model needs to take into account the cost of error when the user misses the target. Often this cost can be abstracted as time required to recover from the error [25].

A simple intuitive approach to model this behavior would be to collect the error-free task completion time in conditions without cost of error, and predict the completion time as the sum of error-free movement time and the cost of error weighted by some observed error rate [27]. However, such a simplification does not take into account that the user's error rate might be a function of the cost which that error incurs.

In this section, we explain in detail our model for task completion time (CT), which takes into account the cost of making an error. Our model is based on a simple assumption that users will change the characteristic of their performance in favor of strategies that maximize their expected utility. In other words, given a penalty represented by a time-based cost of making an error, the users will change their speed and accuracy in order to minimize their task completion time. Although the expected utility theory had a fair share of criticism (e.g., [19]), we assume that the simplicity of target-directed pointing tasks still allows the users to at least have a rough estimate of the optimal utility in such tasks. Additionally, although the probability of an error is not fixed, users should be able to learn to optimize their accuracy [8], which should allow them to use this information in estimating the expected utility.

General-case Target-directed Pointing Task

In a general-case target-directed pointing task, the user attempts to acquire a target of width W at a distance A (Figure 1a). The Shannon formulation of Fitts' law [20] predicts the movement time, MT , for this task to be:

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

However, if the user misses the target on the first attempt, the user often incurs some cost, C . This cost includes the time to recover from the error and to attempt to select the target again. We define the completion time (CT) as the time until the target is successfully selected, whereas movement time (MT) is the time until the first target selection attempt. As such, if the user is successful on their first attempt, CT will equal MT (Figure 1b), and if the user makes an error CT will be the sum of MT and C (Figure 1c). For now we consider the time to reacquire the target successfully as part of the error cost C .

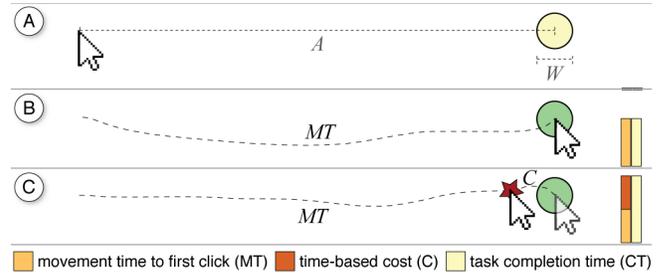


Figure 1. a) A general target-directed pointing task; b) the user selects the target with the first click; c) the user misses the target with the first click, and incurs the cost of error.

Expected Completion Time

In our model, we first calculate the expected completion time ($E[CT]$) of a target acquisition task based on the user's movement time (MT), the user's error rate given that movement time ($P(E)$), and the time-based error cost (C):

$$E[CT] = (1 - P(E)) \cdot MT + P(E) \cdot (MT + C) \quad (2)$$

In order to calculate $P(E)$ we use Wobbrock *et al.*'s [35] error model for pointing in 2 dimensions, which is derived from the Fitts' law equation and predicts the error rate given the user's movement time:

$$P(E) = 1 - \operatorname{erf} \left(\frac{2.066 \cdot \frac{W}{A} \left(\frac{MT e^{-a'}}{2 \frac{b'}{b} - 1} \right)}{\sqrt{2}} \right) \quad (3)$$

Note that a' and b' values in the Equation 3 differ from the a and b values in Equation 1. This is because a' and b' values need to be captured over a range of user speed-accuracy strategies [35]. Therefore, both a and b , and a' and b' have to be computed in order to use our model.

Optimal Expected Completion Time

In order to predict the user's task completion time, we assume that the users will tend towards optimal performance, a hypothesis supported by prior research [8]. At optimal performance the user will be performing at a speed-accuracy tradeoff which minimizes the user's expected completion time.

Therefore, our proposed model for task completion time based on a cost of error C is calculated by minimizing the expected completion time ($E[CT]$):

$$CT = \operatorname{argmin}_{MT} (1 - P(E)) \cdot MT + P(E) \cdot (MT + C) \quad (4)$$

Example For a Specific Error Cost

Our model (Equation 4) abstracts the cost of error as the time-based penalty after an error occurs (C). This time, required for the user to recover from the error and complete the task, is task dependent. Below we show how our model can be applied for a specific error cost.

For demonstration purposes, consider a task where if the user misses the target, a time penalty is applied, the cursor

is frozen in place, and once the penalty expires, the user can attempt to acquire the target again. To model this task, we split C into two components: the time to recover from the error (C_r), and the time to complete the task after the penalty (CT_c). The expected completion time then becomes:

$$E[CT] = (1 - P(E)) \cdot MT + P(E) \cdot (MT + C_r + CT_c) \quad (5)$$

Here, C_r would have some fixed value. CT_c is estimated using Fitts' law, where the width (W) and the position of the target remains the same, but the amplitude (A) changes. We denote this amplitude after an error occurs as A_{err} . We estimate A_{err} by finding the mean of selection endpoints that result in error, and computing the distance from that point to the center of the target (Figure 2). Thus, this starting point is the point with probability equal to $p(z)/2$, where z is the unit-normal deviate, or z -score [29], along the standard normal distribution of selection endpoints.

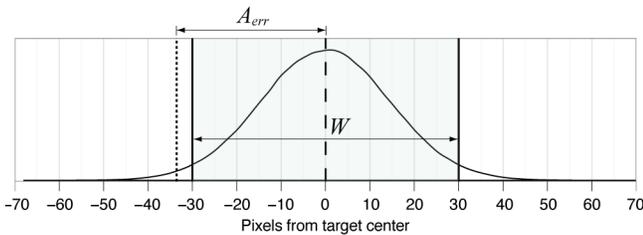


Figure 2. Standard normal distribution of selection endpoints around target of width $W=60$ px in 1 dimension and amplitude after error (A_{err}) in an idealized task when the user's effective movement time results in the 4% error rate.

The movement time and error rate after an error occurs are then calculated by plugging in A_{err} for A :

$$MT = a + b \cdot \log_2 \left(\frac{A_{err}}{W} + 1 \right) \quad (6)$$

$$P(E) = 1 - erf \left(\frac{2.066 \cdot \frac{W}{A_{err}} \left(2^{\frac{MT - a'}{b'}} - 1 \right)}{\sqrt{2}} \right) \quad (7)$$

We then calculate the expected completion time for the task after the error occurs recursively, using Equation 5. In other words, we treat the task after an error occurs as a new target-directed pointing sub-task, where there is a probability that the user will again make an error given the user's effective movement time, and again incur the fixed cost of error C_r if the user misses the target.

With this cost function, we can now use the model to predict task completion times for different time-based costs. Figure 3 shows the relationship between movement MT and $E(CT)$. It can be seen that moving too quickly will cause $E(CT)$ to be high, because of large error rates. Moving too slowly will result in a suboptimal $E(CT)$, as in effect, the behavior overcompensates for the error cost, and results in an error rate that is lower than optimal.

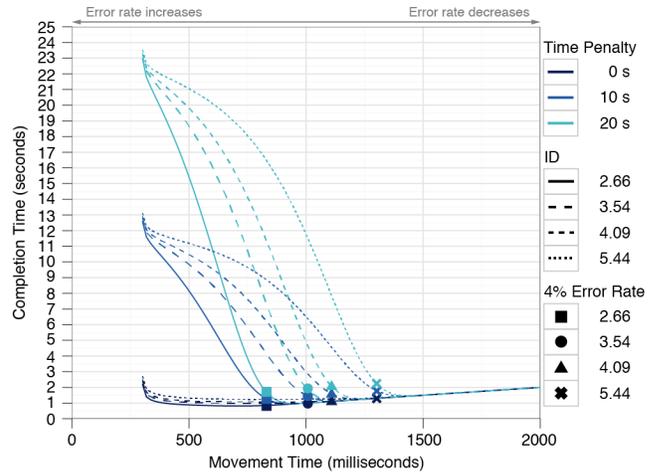


Figure 3. Idealized task completion times per movement time, time-based cost of error, four example indexes of difficulty, and fixed a' and b' values ($a'=300$ ms, $b'=200$ ms/bit).

Figure 4 shows the optimal completion times for different $A \times W$ combinations, based on the time penalty C_r . It can be seen that the optimal completion time curves level off quickly. At first, this might seem surprising because intuitively the users should be performing much slower to avoid the very high penalties. However, Wobbrock *et al.*'s error model [35] suggests that large increases in movement time when the error rate is close to zero results in negligible decrease in error rate. Therefore, the users can move at close to typical movement times even when the costs of errors are extremely high. We believe this to be an interesting insight that results from our model, which we will test in our experiment.

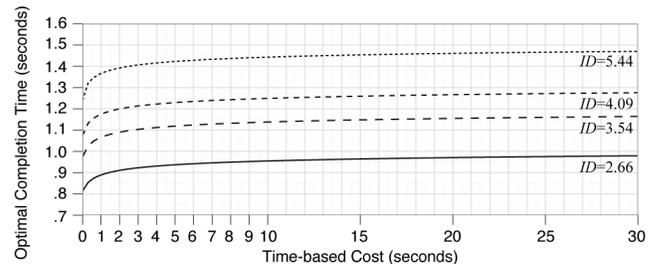


Figure 4. Idealized optimal task completion times per time-based cost of error and four indexes of difficulty, and fixed a' and b' values ($a'=300$ ms, $b'=200$ ms/bit).

EXPERIMENT

The goal of this experiment was to explore whether time-based penalties affect the user performance, and to verify the validity of our cost of error model. We investigate whether the optimal completion time, movement time and error rate computed using our model at different time-based penalties can be used to predict the user's observed performance. The experiment used the task outlined in the above example. That is, if the user made an error, the cursor would freeze for a fixed time-penalty, after which the user was required to successfully acquire the goal target.

Participants

Thirty-six participants (15 female, ages 18-50, median age=35.5) took part in the study. We recruited the participants from an existing participant list. Participants received a \$25 gift card for an online retail store for their participation in the study.

Apparatus

We conducted the experiment on a HP Intel Xeon desktop computer with 3.2GHz dual-processor and 12GB RAM, running Windows 7. The machine was equipped with a 24-inch monitor set to 1920×1200 resolution. All participants used the same HP USB 2-button optical scroll mouse, with its speed set to 7/10 in the mouse control panel and acceleration turned off. The experimental software was implemented using Java 6 SE. The software ran in full-screen mode, with all other nonessential services disabled.

Design

We randomly assigned 18 participants into *short time penalties* group (0, 3.33, and 6.66 seconds) and the other 18 into *long time penalties* group (0, 10, and 20 seconds). We did this instead of asking participants to experience all 5 delays to enable participants to complete more trials per time-based penalty without becoming fatigued.

Participants performed target selections for 3 target sizes (W: 20 px, 60 px, and 100 px) × 3 target distance (A: 160 px, 320 px, and 640 px) × 3 time-based penalties. The study consisted of 3 sections (one for each penalty). Each section consisted of 6 blocks of 15 trials for each of the 9 ID's. We fully counterbalanced the order of the time-based penalties, and randomized the order of ID's within a block. Trials with the same ID were grouped together within a block. Each participant completed 2,430 trials in total.

Task and Procedure

In order to complete a trial, the participants had to successfully click inside a goal target. If they missed the target, they had to continue the trial from where they missed until they correctly selected the target. In a 0 cost condition, users could continue immediately. In the other conditions, the system removed the cursor, displayed a time delay message, and placed the cursor back at the exact same position after the penalty delay. Additionally, the system displayed a reminder about the current time delay at all times in the bottom right corner of the screen. At the beginning of each section, the system required the participant to experience the time-based penalty in a single warm-up trial. In this warm-up trial, the participants had to click inside the first target that appeared on the screen, and then purposely click outside of the next target that appeared to experience what happens when they make an error.

In order to prevent the participants from guessing the next position of the target, and thus impacting their strategy, we used a modified 2-D Fitts' pointing task. At any given time only one circular target was visible on the screen. Once the participant successfully selected the target, a new target would appear in one of eight directions around the previous

target at distance specified by the trial amplitude. The position of the next target was randomized.

After the first and second sections, the system enforced a 5-minute mandatory rest for all participants. At the beginning of the study, an investigator read instructions for the task to the participant, and instructed the participant that the only goal was to complete the entire study as fast as possible. We allowed the participants to perform at error rates that they felt would allow them to accomplish this goal. At the end of the study, the participants filled out a questionnaire asking about their demographics and their strategies for performing the task, to ensure they understood the task.

RESULTS

In this section we present the results of our experiment. Consistent with prior work [29], prior to data analysis, we removed spatial outliers from the data where the amplitude of movement was less than half the nominal movement amplitude, or where the selection endpoint was more than twice the target width from the target center. This removed 64 (0.07%) trials. In order to calculate the effective amplitude (A_e) and width (W_e) of the target, we calculated bivariate deviation of click points in 2 dimensions ($SD_{x,y}$) as a spread of hits around the target centroid [36].

We analyzed data for *short* and *long time delay* groups separately. We analyzed parametric participant data, such as movement time, with one-way repeated measure ANOVAs. To ensure Sphericity of data, we performed Mauchly's tests, and used the Greenhouse-Geisser correction when Sphericity was violated. We performed pair-wise comparison using paired t-tests with Bonferroni correction. We analyzed non-parametric data, such as error rates, using Friedman tests and performed pair-wise comparison using Wilcoxon tests.

Effects of Time-based Penalty on Task Performance

In this section, we investigate the effects of time-based penalty on the participants' task completion time, movement time to the first click, and error rate.

Completion Time

Figure 5 shows mean observed completion times for the two time penalties groups. The time-based penalty had a significant effect on observed movement time in the *short time penalties* group ($F_{2,34}=31.86, p<.001, \eta_p^2=.65$) and the *long time penalties* group ($F_{2,34}=38.24, p<.001, \eta_p^2=.69$). In the *short time penalties* group, the participants' mean completion times were 944 ms for 0, 1046 ms for 3.33, and 1096 ms for 6.66 seconds penalty. All were significantly different (all $p<.05$). In the *long time penalties* group, the participants on average completed trials faster when the penalty was 0 second (mean=969 ms) than when the penalty was 10 (mean=1133 ms; $p<.001$) and 20 seconds (mean=1178 ms; $p<.001$). But our tests did not find significant difference between 10 and 20 seconds penalties ($p=.29$). This is consistent with our model that predicts that the cost of error will have less impact at higher cost values.

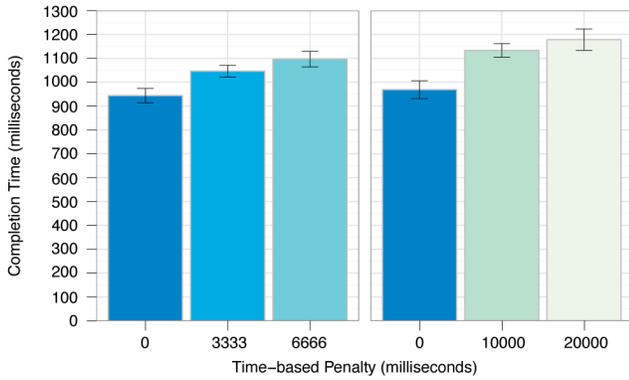


Figure 5. Mean observed completion times per time-based penalty for short time penalty (left) and long time penalty (right) groups. Error bars represent 95% confidence interval.

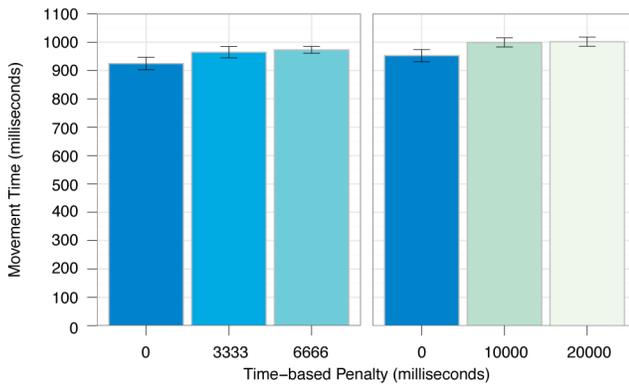


Figure 6. Mean observed movement times per time-based penalty for short time penalty (left) and long time penalty (right) groups. Error bars represent 95% confidence interval.

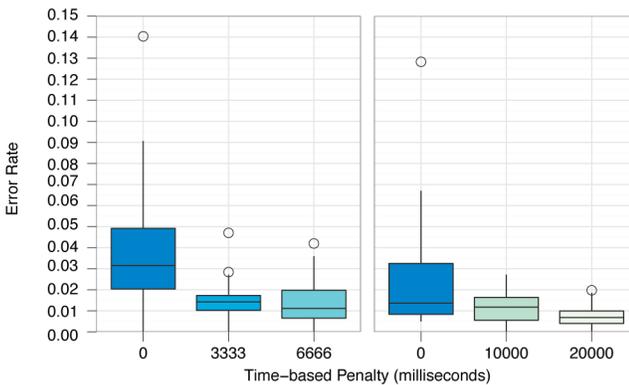


Figure 7. Median observed and optimal error rates per time-based penalty for short time penalty (left) and long time penalty (right) groups.

Movement Time

The time-based penalty had a significant effect on observed movement time before the first target selection in both short time penalty ($F_{2,34}=8.82, p<.001, \eta_p^2=.34$) and long time penalty ($F_{2,34}=10.34, p<.001, \eta_p^2=.38$) groups. The participants were faster when the penalty was 0 seconds (mean=924 ms) than when the penalty was 3.33 seconds (mean=964 ms; $p=.0577$) and 6.66 seconds (mean=973 ms; $p=.0016$). Our tests did not find significant difference

between 3.33 and 6.66 seconds penalties ($p=1$). Similarly, participants performed faster when the penalty was 0 seconds (mean=952 ms) than when it was 10 seconds (mean=999ms; $p=.0085$) and 20 seconds (mean=1002 ms; $p=.0054$), but our tests did not find significant difference between 10 and 20 second penalties ($p=1$) (Figure 6). However, if the movement times, and thus error rates, were identical across different time-based costs, the completion times would show linear increase. Instead, the completion time data appears to follow the curve estimated by our model. The perfectly increasing order of movement time as cost increases also provides evidence that participants are responding to the level of cost, and not just its existence.

Error Rate

The time-based penalty also had a significant effect on the error rate in the short time penalty ($\chi^2_{(2)}=12.11, p=.0023$) and long time penalty ($\chi^2_{(2)}=20.69, p<.001$) groups (Figure 7). Participants in the short time penalty group made more errors in the 0 second (median=0.0316) condition than in the 3.33 second (median=0.0142; $p=.002, r=.57$) and 6.66 second (median=0.0111; $p=.0038, r=.54$) conditions, but our tests did not find a significant difference between 3.33 second and 6.66 second penalties ($p=.71, r=.20$). Similarly in the long time penalty group, the participants made more errors when the penalty was 0 seconds (median=0.0137) than when it was 10 seconds (median=0.0118; $p=.0091, r=.50$) and 20 seconds (median=0.0068; $p<.001, r=.62$). Our tests again did not find significant difference between 10 and 20 second penalties ($p=.24, r=.28$). It should be noted that our model predicts that this decrease in error rate will become small as we increase the cost of error (e.g. between 10 and 20 seconds) because the participants already have to perform at close to 0% error rates to minimize their completion time.

Completion Time Model Verification

Having confirmed that participants do adjust their strategy based on time penalties, we now validate our model's calculation of $P(E)$ (Equation 3) and $E[CT]$ (Equation 5).

Predicting Error Rates Based on MT

For each participant, we first applied Crossman's correction [5] in order to align the data across different time-based penalty groups. We then computed the participant's effective movement time a and b values by fitting the observed movement time to the first target selection to the effective indexes of difficulty from our task.

However, the error model equation (Equation 3) [35] requires a' and b' values different from the a and b values above. Therefore, we calculated the a' and b' values for each participant by using their observed movement time and error rate. We then derived the a' and b' values by minimizing the mean squared error between the observed and model predicted error rate. Note that the predicted error rate computation requires an effective movement time term (Equation 3). We computed this effective movement time using the fixed a and b values for each participant.

Figure 8 shows $N=54$ error rate points for each $A \times W \times \text{penalty group} \times \text{time penalty}$ plotted as (*predicted, observed*) ordered pairs averaged over all participants. Linear regression results in $y=.0072+.87 \cdot x$, ($R^2=.55$) and the model fit was significant ($F_{1,52}=63.89$, $p<.001$). Mean difference between the predicted and observed errors was 0.008. Note, however, that although the linear fit has a near-zero intercept and near-unity slope, R^2 is low. One possible explanation is that the observed error rates in our experiment take values only between 0 and 0.05, but the original model was derived from observed error rates ranging from 0.0 to 0.7.

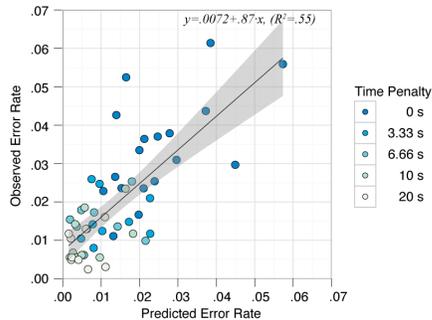


Figure 8. Predicted vs. observed error rate. The shaded area represents 95% confidence interval.

Predicting Expected Completion Times Based on MT

In this section we verify our calculation of expected completion time (Equation 5). We calculated the expected completion time using the observed *MT* from the participants (Figure 6). The error rates were predicted by the error model (Equation 3) given the observed *MT* and a' and b' values we derived in the previous section.

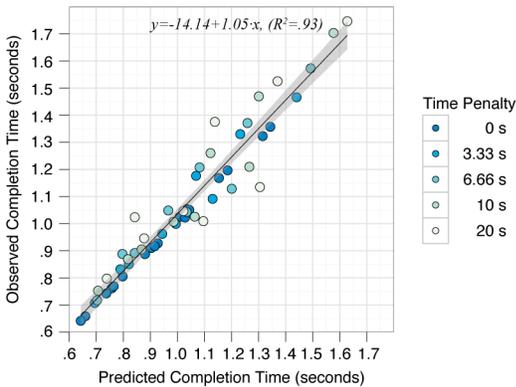


Figure 9. *MT*-based prediction vs. observed completion times. The shaded area represents 95% confidence interval.

Figure 9 shows $N=54$ completion time points for each $A \times W \times \text{penalty group} \times \text{time penalty}$ plotted as (*predicted, observed*) ordered pairs averaged over all participants. Linear regression results in $y=-14.14+1.05 \cdot x$ ($R^2=.93$) and the model fit was significant ($F_{1,52}=709$, $p<.001$). Mean difference between the predicted and observed completion times was 54.16 milliseconds. The near-zero intercept and near-unity slope suggest that, given the participants'

movement time, the expected completion time equation predicts the observed completion times well.

Predicting User Performance Based on Error Cost

In this section, we explore if the optimal performance computed with our model can be used to estimate the participants' performance at different time-based penalties.

Optimal Completion Time

We first compare the optimal completion time calculated using our model and the observed completion time. Figure 10 shows the data for $N=54$ points for each $A \times W \times \text{penalty group} \times \text{time penalty}$ plotted as (*optimal, observed*) ordered pairs averaged over all participants. This resulted in a fit with equation $y=-62.29+1.18 \cdot x$, ($R^2=.93$), and the fit was significant ($F_{1,52}=655.30$, $p<.001$). The near-zero intercept and near-unity slope suggest that our completion time model predicts the observed completion time well.

The mean observed completion times averaged over all combinations of $A \times W$ were slower than optimal. This indicates participants will approach, but not actually achieve the theoretically calculated optimal time. The relationship between optimal and actual completion times can be described as:

$$CT_{actual} = -62.29 + 1.18 \cdot CT_{optimal} \tag{8}$$

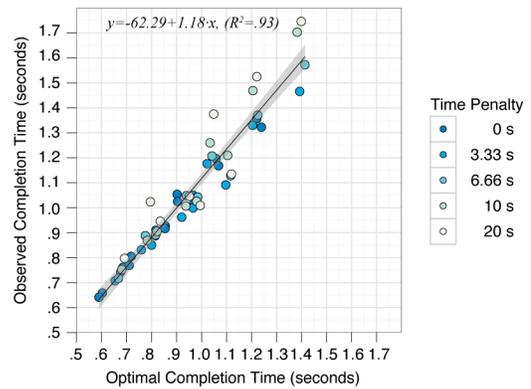


Figure 10. Optimal vs. observed completion time. The shaded area represents 95% confidence interval.

Optimal Movement Time

Figure 11 left shows a comparison between the optimal movement time calculated using our completion time model and the observed movement time for $N=54$ points for each $A \times W \times \text{penalty group} \times \text{time penalty}$ plotted as (*optimal, observed*) ordered pairs averaged over all participants. The fit between the two movement times resulted in equation $y=150.52+.91 \cdot x$, ($R^2=.85$). Although the fit has a near-unity slope the intercept is very high. Upon further investigation of the data we found that this was due to movement times in the 0 second condition being slower than the optimal movement time, by an average of 17.55%. We then compared optimal movement time with the observed movement time for all time-based penalties except 0 second penalty ($N=36$) (Figure 11 right). This resulted in a better fit ($y=24.71+.99 \cdot x$, ($R^2=.98$)).

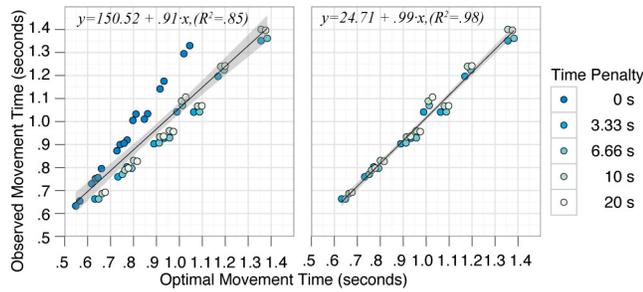


Figure 11. Optimal vs. observed movement time for all time-based penalties (left) and with 0 second penalty data points removed (right). The shaded area represents 95% confidence interval.

Optimal Error Rate

Similarly, we compare the observed error rate with the predicted optimal error rate (Figure 12). The fit between the observed and optimal error rates resulted in equation $y = .0118 + .1156 \cdot x$, ($R^2 = .73$), and the fit was significant ($F_{1,52} = 142.40$, $p < .001$). However, the slope in the equation indicates that the optimal error rate provides a poor prediction of the participants' observed error rate. One possible explanation again is that the participants performed sub-optimally in the 0 second penalty condition and did not make nearly as many errors as it would be optimal. Therefore, we compare the observed error rate with the optimal error rate for all except 0 second time penalty. This fit of $N = 36$ points resulted in equation $y = .0078 + 1.2294 \cdot x$, ($R^2 = .28$), and the fit was again significant ($F_{1,34} = 13.19$, $p < .001$). Although the intercept became closer to 0 and the slope came closer to unity, the R^2 lowered.

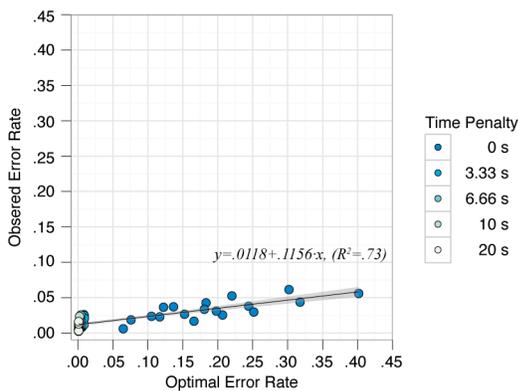


Figure 12. Optimal vs. observed error rates. The shaded area represents 95% confidence interval.

DISCUSSION

Our results show that users change the characteristics of their target-directed pointing given different time-based penalties. These changes are more pronounced for the smaller penalties (0-6 seconds) and have the most effect on the task completion time. This is in line with the idealized curves generated using our model which suggest that differences in performances get smaller as time penalties increase. Nevertheless, this shows that the 0 second penalty condition often used in target-directed pointing tasks does

not generalize to tasks involving costs of making an error. Therefore, in order to calculate the total completion time for a task where an error results in time-based penalty, a model such as ours is needed.

The results also show that the participants change their performance according to the expected completion time function (Equation 5). We show that the expected completion times predict the observed completion times well given the user's movement time to first target selection. In addition, we show that Wobbrock *et al.*'s error model for pointing [35] is robust enough to allow us to interpolate the a' and b' model parameters from a smaller range of error rate values (0%-7%).

Our results show that optimal completion times, calculated using our model, predict the task completion times well. Our model can then be used to interpolate task completion times for time-based penalties other than those tested. In our work, parameters for the movement time function are derived based on Fitts Law, while parameters of the error rate function are derived based on Wobbrock *et al.*'s error model [35]. These values were all derived independent to our model, and then the fit of our model, using these values, was tested. Therefore, our model allows for predicting task completion times for tasks that involve the cost of making an error; something existing target-pointing models do not.

The model can additionally be used to predict other parameters of the task, such as the movement time to the first selection. However, the model best predicts movement times for costs other than the 0 second cost. This is due to the participants performing significantly slower than optimal in the 0 second condition. One reason for this could be that the participants chose from many possible strategies [16]. For example, participants could have been too risk-averse [25] or they tended towards "satisfactory" rather than optimal performance [11]. It is also possible that participants considered energy expenditure [8] as well when optimizing their performance. This energy expenditure dimension was probably most pronounced in the 0 second penalty conditions because participants were not able to offer constant performance at the optimal movement time.

The optimal error rates calculated using our model did not predict the observed error rates as well as the other task parameters. This was especially true for the observed error rates in the 0 second condition. This could also be due to the slower movement times in this condition. However, even in conditions with time-based penalty greater than 0, the model does not provide a strong fit. One possible explanation for this is that the spread of error rates for conditions with time-based penalty greater than 0 seconds we tested was too narrow (0%-3%). Also, it is possible that the sample size (the number of trial repetitions) was not large enough to average out the noise in the error rate data.

Nevertheless, our model provides a framework for predicting the completion times for tasks that include the

cost of error. As such, our model could use other target-directed pointing movement time and error rate models to alleviate some of the limitations we discussed above. For example, a time-accuracy model for movement time that accounts for the energy expenditure [14] could be used instead of the Shannon formulation of Fitts' law we used.

While some of the differences between completion times predicted by our model are fairly small, our model is the first to provide an explanation and prediction of changes in user behavior given the cost of error when the user misses the target. Additionally, even such small timesavings can have an impact on the users' behavior [12]. Our model, therefore, could inform design decisions on whether reducing the cost of errors of different target-directed pointing tasks in a user interfaces will result in significant performance advantages.

The optimal performance calculated using our model has implications for other target-directed pointing tasks metrics. For example, past research referred to the ideal target utilization as the utilization that corresponds to the 4% error rate [40]. However, our results show that the target utilization varies for different levels of time-based penalty. What might be considered underutilization of a target for one time-based penalty might be considered ideal utilization or overutilization of the target for a higher time penalty. Therefore, the ideal target utilization could be defined as the utilization that corresponds to optimal error rate for the given time penalty. The error cost might have impact on other target-directed pointing metrics as well, such as throughput [21, 40], which require further investigation.

CONCLUSION AND FUTURE WORK

In this paper we presented our model which predicts the completion times for target-directed pointing tasks given the time-based cost of error. We described an experiment in which we computed the task parameters for our model, including the parameters used to derive the pointing error model we use in our model. The results showed that our model accurately predicts user performance in target-directed pointing tasks that involve the cost of error when the user misses the target. We then provided a discussion about some of the limitations of our model and how those limitations could be addressed in the future. Additionally, our work opens up a number of interesting opportunities for future research.

We briefly discussed the impact our model could have on some metrics of target-directed pointing. However, future work should perform a more thorough investigation of the effects that cost of errors has on other task metrics, such as for example throughput [21, 40].

In our work we examined the effect of cost of error on a target-directed pointing task. It would be interesting to explore how our model contributes to existing cognitive models for tasks, such as for example in text entry tasks [3,

22, 39]. Additionally, future work should explore the effects of cost of error on other common tasks the users perform in user interfaces. One such task is the steering task where the impact of the cost of error might be even greater than in the pointing task.

Finally, our model abstracts the cost of error as the time required to recover from the error. However, there are different types of costs that the user might incur when interacting with user interfaces. The mental cost and the physical cost of making an error, *i.e.*, the mental and physical demand required to recover from the error, are two such cost types that would be interesting to explore in future research.

In summary, we believe that taking into account the cost of errors during interactive tasks is an interesting and important area, and believe that our research can serve as groundwork for future investigations on this topic.

ACKNOWLEDGMENTS

The authors wish to thank Jacob O. Wobbrock for valuable feedback about the error model based on the Fitts' law we used in our model.

REFERENCES

1. Adam, J. J. (1992) The effects of objectives and constraints on motor control strategy in reciprocal aiming movements. *J. Motor Behav.* 24, 173-185.
2. Al-Imam, E. and Lank, E. (2006) Biasing response in Fitts' Law tasks. In *CHI EA '06*. ACM, 460-465.
3. Arif, A. S. and Stuerzlinger, W. (2010) Predicting the cost of error correction in character-based text entry technologies. In *Proc. CHI '10*. ACM, 5-14.
4. Card, S. K., Moran, T. P. and Newell, A. (1983) *The Psychology of Human-Computer Interaction*. Lawrence Erlbaum, Hillsdale.
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.). The M.R.C. and D.I.S.R. Joint Committee on Individual Efficiency in Industry.
6. Dean, M., Wu, S.-W., & Maloney, L. T. (2007) Trading off speed and accuracy in rapid, goal-directed movements. *J. of Vision*, 7 (5), 1-12
7. Douglas, S. A., Kirkpatrick, A. E., and MacKenzie, I. S. (1999) Testing pointing device performance and user assessment with the ISO 9241, Part 9 standard. In *Proc. CHI '99*. ACM, 215-222.
8. Elliott, D., Hansen, S., Mendoza, J., and Tremblay, L. (2004) Learning to optimize speed, accuracy, and energy expenditure: a framework for understanding speed-accuracy relations in goal-directed aiming. *J. Motor Behav.* 36 (3), 339-351.
9. 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.

10. 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.
11. Gigerenzer, G. and Selten, R. (2002). *Bounded Rationality: The Adaptive Toolbox*. MIT Press.
12. Gray, W. D. and Boehm-Davis, D. A. (2000) Milliseconds matter: an introduction to microstrategies and to their use in describing and predicting interactive behavior. *J. of Exp. Psychology: A*, 6 (4), 322-335.
13. Gray, W. D., Sims, C. R., Fu, W., and Schoelles, M. J. (2006) The soft constraints hypothesis: a rational analysis approach to resource allocation for interactive behavior. *Psychological Review*, 113 (3), 461-482.
14. Guiard, Y., Olafsdottir, H. B., and Perrault, S. T. (2011) Fitts' law as an explicit time/error trade-off. In *Proc. CHI '11*. ACM, 1619-1628.
15. Hancock, P. A. and Newell, K. M. (1985) The movement speed-accuracy relationship in space-time. In *Motor Behavior: Programming, Control, and Acquisition*, Heuer, H., Kleinbeck, U., and Schmidt, K.-H. (eds). Springer-Verlag, 153-188.
16. Howes, A., Lewis, R. L., and Vera, A. (2009) Rational adaptation under task and processing constraints: Implications for testing theories of cognition and action. *Psychological Review*, 116, 4, 717-751.
17. Hudson, T. E., Maloney, L. T., and Landy, M. S. (2007) Movement Planning With Probabilistic Target Information. *J. Neurophys* 98, 3034-3046.
18. Hudson, T. E., Maloney, L. T., and Landy, M. S. (2008) Optimal Compensation for Temporal Uncertainty in Movement Planning. *PLoS Comput. Biol.* 4(7).
19. Kahneman, D. and Tversky, A. (1979) "Prospect Theory: An Analysis of Decision under Risk", *Econometrica*, XLVII, 263-291.
20. MacKenzie, I. S. (1992) Fitts' law as a research and design tool in human-computer interaction. *Human Computer Interaction* 7 (1), 91-139.
21. MacKenzie, I. S. and Isokoski, P. (2008) Fitts' throughput and the speed-accuracy tradeoff. In *Proc. CHI '08*. ACM, 1633-1636.
22. MacKenzie, I. S. and Zhang, S. X. (1999) The design and evaluation of a high-performance soft keyboard. In *Proc. CHI '99*. ACM, 25-31.
23. 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, 340-370.
24. 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*, Jeannerod, M. (ed). Lawrence Erlbaum, 173-226.
25. Munichor, N., Erev, I., and Lotem, A. (2006) Risk attitude in small timesaving decisions. *J. of Exp. Psychology: Applied*, 12 (3), 129-141.
26. Plamondon, R. and Alimi, A. M. (1997) Speed/accuracy trade-offs in target-directed movements. *Behav. Brain Sci.* 20 (2), 279-349.
27. Reason, J. (1990), *Human Error*. Cambridge University Press, Manchester.
28. Schoemaker, P. J. H. (1982). The Expected Utility Model: Its Variants, Purposes, Evidence and Limitations. *J. of Economic Literature* 20, 529-563.
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. J. Hum.-Comput. Stud.* 61 (6), 751-789.
30. Soukoreff, R. W. and MacKenzie, I. S. (2009) An informatic rationale for the speed-accuracy tradeoff. In *Proc. SMC'2009*. IEEE, 2890-2896.
31. Trommershäuser, J., Maloney, L. T., and Landy, M. S. (2003) Statistical decision theory and rapid, goal-directed movements. *J. of the Optical Soc.*, 1419-1433.
32. Trommershäuser, J., Maloney, L. T., and Landy M. S. (2008) The expected utility of movement. *Neuroeconomics*, Academic Press, 95-111.
33. Welford, A. T. (1968) *Fundamentals of Skill*. Methuen, London.
34. Wobbrock, J. O., Cutrell, E., Harada, S., and MacKenzie, I. (2008) An error model for pointing based on Fitts' law. In *Proc. CHI '08*. ACM, 1613-1622.
35. Wobbrock, J. O., Jansen, A., and Shinohara, K. (2011) Modeling and predicting pointing errors in two dimensions. In *Proc. CHI '11*. ACM, 1653-1656.
36. Wobbrock, J. O., Shinohara, K., and Jansen, A. (2011) The effects of task dimensionality, endpoint deviation, throughput calculation, and experiment design on pointing measures and models. In *Proc. CHI '11*. ACM, 1639-1648.
37. Woodworth, R. S. (1899) The accuracy of voluntary movement. *Psychological Review, Monograph Supplements* 3 (2, Whole No. 13), 1-114.
38. Young, S. J., Pratt, J., Chau, T. (2009) Target-Directed Movements at a Comfortable Pace: Movement Duration and Fitts's Law. *J. Motor Behav.* 41, 339-346.
39. Zhai, S., Hunter, M., and Smith, B. A. (2000) The metropolis keyboard - an exploration of quantitative techniques for virtual keyboard design. In *Proc. UIST '00*. ACM, 119-128.
40. 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.