Modeling and Predicting Pointing Errors in Two Dimensions

Jacob O. Wobbrock, Alex Jansen, Kristen Shinohara
The Information School
DUB Group
University of Washington
Seattle, WA 98195 USA
{wobbrock, ajansen7, kshino}@uw.edu

ABSTRACT
Recently, Wobbrock *et al.* (2008) derived a predictive model of pointing accuracy to complement Fitts' law's predictive model of pointing speed. However, their model was based on one-dimensional (1-D) horizontal movement, while applications of such a model require two dimensions (2-D). In this paper, the pointing error model is investigated for 2-D pointing in a study of 21 participants performing a time-matching task on the ISO 9241-9 ring-of-circles layout. Results show that the pointing error model holds well in 2-D. If univariate endpoint deviation (SD_x) is used, regressing on N=72 observed *vs.* predicted error rate points yields R^2=.953. If bivariate endpoint deviation ($SD_{x,y}$) is used, regression yields R^2=.936. For both univariate and bivariate models, the magnitudes of observed and predicted error rates are comparable.

Author Keywords: Pointing error model, Fitts' law, metronome, movement time, error prediction, error rates.

ACM Classification Keywords: 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, Measurement, Theory.

INTRODUCTION

Fitts' law [3,6] predicts movement time (MT) in rapid aimed pointing tasks with the following formulae:

$$MT = a + b \cdot ID, \quad (1)$$

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

ID is the task index of difficulty measured in *bits*. Task parameters are target distance A and size W, and a and b are empirical regression coefficients.

Recently, Wobbrock *et al.* [13] derived a predictive model of pointing accuracy (Eq. 3) to complement Fitts' law's predictive model of pointing speed. The independent variables in the model include target distance (A) and size (W), but now instead of predicting MT, time becomes the *actual* time taken to reach the target, MT_e. Thus, movement time becomes an *independent* variable in the accuracy model instead of the dependent result of Fitts' law. Eq. 3 uses the error function (*erf*); see footnote 1 in [13].

$$P_{error} = 1 - erf\left\{\frac{1}{A\sqrt{2}}\left[2.066 \cdot W\left(2^{\frac{MT_e - a'}{b'}} - 1\right)\right]\right\} \quad (3)$$

It is when a user's movement time may not be predictable by Fitts' law that the above pointing error model is potentially most useful. Haste, tentativeness, or extra care may cause users to point at speeds not predicted by Fitts' law. For example, in computer games, targets often appear for short durations, forcing players to rush. Another example is predicting text entry error rates on stylus keyboards for users that deliberately slow down or speed up. Yet another example is in safety-critical interfaces when trying to make controls big enough in light of space constraints to ensure a given error rate. In all instances, a pointing error model is required to make quantitative performance predictions. A current limitation, however, is that Eq. 3 was based on horizontal pointing to one-dimensional (1-D) vertical ribbon targets, but applications of a pointing error model require two dimensions (2-D).

In this paper, we investigate the accuracy of the pointing error model using the 2-D multidirectional ring-of-circles arrangement from the ISO 9241-9 standard [4,10]. As in Wobbrock *et al.*'s 1-D study [13], we use a time-matching metronome task to control MT_e as an independent variable. The metronome is paced individually for each subject by first establishing that subject's Fitts' law model.

Our findings show that the pointing error model holds well in 2-D. If univariate endpoint deviation (SD_x) is used while fitting Fitts' law, regressing on N=72 observed (y) *vs.* predicted (x) error rate points yields R^2=.953 with an equation of $y = .026 + 1.045x$. Using bivariate endpoint deviation ($SD_{x,y}$) yields R^2=.936 with an equation of $y = -.002 + 0.916x$. (Having R^2=1.00 and y=x would be a "perfect model.") Thus, both outcomes show well-fit models and comparable magnitudes of observed and predicted error rates.

REVIEW OF THE POINTING ERROR MODEL

Here, we do not re-derive the pointing error model of Eq. 3, but instead highlight aspects relevant to our inquiry. Readers are directed to prior work [13] for the derivation.

Important to the derivation of the pointing error model is the observation that endpoints in rapid aimed movements follow a normal distribution,[1] and that the entropy therein is $\log_2 \sigma\sqrt{2\pi e}$, where σ is the standard deviation of the distribution. As Welford [12] (pp. 147-148) and MacKenzie [6] explain, this results in about 96% of endpoints falling within the distribution, and about 4% falling outside. When a ~4% error rate occurs in an experiment, $\log_2 W$ accurately reflects the information in the distribution. But when subjects point with a higher error rate, their "effective" target width W_e is greater than W. When they point with an error rate less than 4%, $W_e < W$. Thus, W_e enables a *post hoc* adjustment to a 4% error rate. This is the basis for Crossman's speed-accuracy correction [1], which uses W_e instead of W in Eq. 2 and allows a "fast but careless" subject to be compared to a "slow and careful" subject. When fitting Fitts' law models to subjects, the effective index of difficulty (ID_e) in Eq. 4 replaces the nominal ID from Eq. 2, effecting Crossman's correction.

$$ID_e = \log_2\left(A_e/W_e + 1\right), \tag{4}$$

$$W_e = 4.133 \cdot SD_x \tag{5}$$

In Eq. 4, A_e reflects the mean distance of actual movements. In Eq. 5, SD_x is the univariate standard deviation of endpoint x-coordinates for an $A \times W$ condition whose data have been rotated to a reference axis (*e.g.*, 0°).

The pointing error model assumes that if a subject points at the speed with which Fitts' law predicts they should (*i.e.*, $MT_e = MT$), they will point with a ~4% error rate.[2] If they point faster than Fitts' law predicts they should (*i.e.*, $MT_e < MT$), the error rate will rise above 4%. If they point slower (*i.e.*, $MT_e > MT$), the error rate will drop below 4%.

In Eq. 3, the regression coefficients a' and b' are decorated with primes to indicate they are not from a traditional Fitts' law study but instead come from fitting Fitts' law to, ideally, a range of $MT_e \neq MT$. Put another way, a' and b' should be built upon movements spanning the speeds for which one intends to predict error rates.

Wobbrock *et al.*'s [13] investigation of the pointing error model in 1-D resulted in a model fit of R^2=.959 for N=90 observed (y) *vs.* predicted (x) points with a regression equation of $y = .007 + 0.958x$, similar to our results here.

A NOTE ON ENDPOINT DEVIATION IN 2-D

As described above, the pointing error model relies on Fitts' law regression coefficients a' and b', which arise from fitting a line to a subject's (ID_e, MT_e) points; ID_e relies on W_e, which relies on SD_x, the spread of hits. In both 1-D and 2-D, this spread of hits can be defined as a deviation of coordinates around their centroid [2,5]. A deviation of 1-D coordinates around a centroid reduces to the standard deviation formula for scalars. However, in 2-D, this reduction does not occur [14]; accordingly, we have the option of either using SD_x and ignoring deviation orthogonal to the task axis, or of using bivariate deviation $SD_{x,y}$ and capturing deviation in both dimensions. For completeness, we explore error predictions using both univariate (SD_x) and bivariate ($SD_{x,y}$) endpoint deviation as the basis for W_e and ID_e. See [14] for further discussion.

METHOD

Subjects and Apparatus

Twenty-one subjects participated in our study, of which a third were female. All subjects were right-handed. Average age was 29.3 years (SD=6.9). Subjects were run on a 21" Samsung SyncMaster 214T flat panel monitor set to 1600×1200 resolution. The computer had a Xeon CPU running Windows 7 at 3 GHz with 2 GB RAM. The input device was a Logitech optical mouse. We built a full-screen program in C♯ called *FittsStudy* [14] to facilitate the study.

Procedure

Our study consisted of two parts, both of which used the ISO 9241-9 ring-of-circles target arrangement [4,10] with 23 targets per ring. The first three of these targets were logged as practice and ignored during analysis, leaving 20 test trials per ring. A trial began immediately following the click of the previous trial and ended with a single click, regardless of whether that click hit the target. Misses were accompanied by a red flash and an audible ding.

Part 1. Subjects' individual Fitts' law models were built. FittsStudy was configured to administer 18 $A \times W$ conditions defined by 3 levels of A {256, 384, 512 pixels} and 6 levels of W {8, 16, 32, 64, 96, 128 pixels}, yielding 13 unique IDs ranging from 1.59 – 6.02 bits. Subjects' Fitts' law models were used to parameterize the metronome in part 2.

Part 2. A metronome-based time-matching study was run.[3] In this part, the same 18 $A \times W$ conditions were used, but now they were crossed with 4 levels of $MT_{\%}$ {0.6, 0.8, 1.0, 1.2}, a factor that, when multiplied by a subject's Fitts' law-predicted movement time (MT), gave the raw speed of the metronome as MT_{ms} (Eq. 6). By manipulating $MT_{\%}$ instead of raw MT_{ms}, subject-specific speed differences were accommodated.

$$MT_{ms} = MT_{\%} \times MT \tag{6}$$

The metronome was both visual and auditory, with the current circle target progressively flood-filling from the center outward and an audible "tick" playing the moment the fill completed, which was also the moment the subjects were supposed to click. After this moment, the flood fill disappeared and began to grow again.

[1] In 2-D, deviations in both x and y are normally distributed [8], but often with different extents.

[2] One can verify this by substituting $a + b \log_2(A/W + 1)$ for MT_e in Eq. 3, which gives $P_{error} = 1 - erf(2.066/\sqrt{2}) \approx .039$, about 4%.

[3] Metronomes are, of course, not part of everyday computer use, but they have been employed in motor psychology to control movement time [9]. A metronome is *not* required for the pointing error model, only the actual movement time MT_e (see Eq. 3).

In part 1, 21 subjects completed 18 $A \times W$ conditions each comprising 20 test trials for a total of 7560 trials. In part 2, the same 21 subjects completed 72 $MT_{\%} \times A \times W$ conditions each comprising 20 test trials for a total of 30,240 trials.

Analysis

When fitting Fitts' law to a subject, each $A \times W$ ring of 20 test trials resulted in one (ID_e, MT_e) point, giving 18 points for regression per subject in part 1 and 72 points in part 2. As stated above, part 1's *a* and *b* coefficients were used to set the metronome, while part 2's a' and b' coefficients were used in Eq. 3 for error rate prediction. Crossman's speed-accuracy correction [1] was used for each subject in each condition. All trials within an $A \times W$ or $MT_{\%} \times A \times W$ condition were normalized to horizontal (0°) before endpoint deviations were computed as SD_x and $SD_{x,y}$.

RESULTS

Adjustment of Data

For part 1, spatial outliers were removed. They were defined from prior work [7] as errors whose effective distances were less than half their nominal distances, or whose endpoints fell more than twice their target widths from their target centers. In all, 2 spatial outliers were removed from part 1 (0.03%). For part 2, spatial outliers were kept, but temporal outliers were removed, defined as movements whose actual durations were shorter than 75% or longer than 125% of the metronome interval. In all, 340 temporal outliers were removed from part 2 (1.12%).

Part 1, Fitts' Law Models ($A \times W$)

In part 1, the average movement time was 761 ms with an average error rate of 4.3%, close to the desired 4%. Using a mean-of-means throughput calculation [10], throughputs were 4.91 bits/s and 4.49 bits/s with univariate (SD_x) and bivariate ($SD_{x,y}$) endpoint deviation, respectively. Average Fitts' law fits were R^2=.951 and R^2=.962 for N=18 points, respectively. Thus, the *a* and *b* model values elicited for each subject in part 1 for pacing the metronome in part 2 were trustworthy.

Part 2, Pointing Error Models ($MT_{\%} \times A \times W$)

Overall Results and Fit of Fitts' Law

In part 2, the average movement time MT_e and average metronome time MT_{ms} should be close. Indeed they were, at 710 ms and 696 ms, respectively, a ratio of 1.02. The average overall error rate was 18.9%. For $MT_{\%}$=1.0, the average error rate was 6.07%, near the desired 4% rate. Throughputs were 4.83 bits/s and 4.39 bits/s using SD_x and $SD_{x,y}$, respectively. Average Fitts' law fits for N=72 points were R^2=.944 and R^2=.948, respectively. Thus, the a' and b' model values elicited for each subject and used in Eq. 3 for error rate prediction were trustworthy.

Overall Fit of the Pointing Error Model

Overall, predicted error rates were 15.6% and 20.9% using univariate (SD_x) and bivariate ($SD_{x,y}$) endpoint deviation, respectively, spanning the actual error rate of 18.9%. Figure 1 shows the N=72 points for each combination of $MT_{\%} \times A \times W$ used in the study plotted as observed (y) *vs.* predicted (x) error rates. Both SD_x and $SD_{x,y}$ result in good fits for the pointing error model with high R^2 values and regression equations near y=x. A Wilcoxon signed-rank test on absolute differences between predicted and observed error rates using each deviation scheme shows no significant difference between the SD_x and $SD_{x,y}$ models (p=.19), indicating they are comparable in fitting these data.

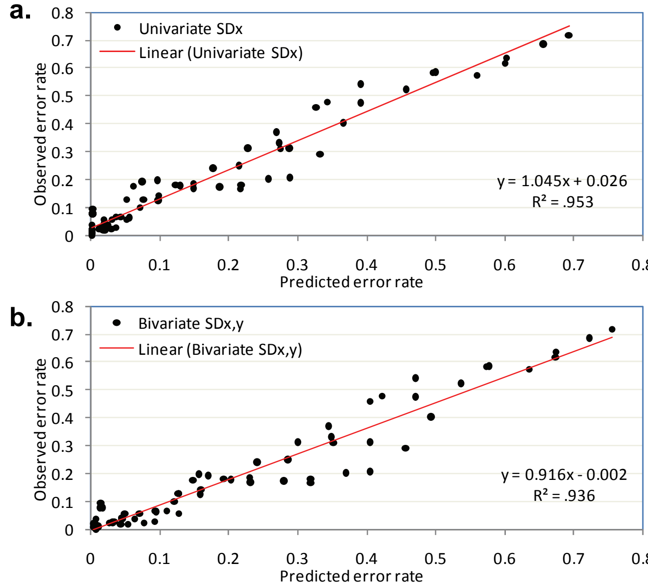

Figure 1. Error model observed (y) *vs.* predicted (x) points for our study's 72 $MT_{\%} \times A \times W$ conditions using **(a)** univariate (SD_x) and **(b)** bivariate ($SD_{x,y}$) endpoint deviation.

Effect of $MT_{\%}$ on Observed Error Rate

We found the relationship between $MT_{\%}$ and observed error rate O_{error} to be logarithmic, best fit by the equation $\ln(O_{error}) = 2.594 - 5.557 \cdot MT_{\%}$ (R^2=.807, N=72 points). Unsurprisingly, as $MT_{\%}$ increased, O_{error} significantly decreased ($\chi^2_{(3,N=72)}$=58.88, p<.0001).

Effect of ID, A and W on Observed Error Rate

In accord with Wobbrock *et al.*'s [13] 1-D study, A had no significant effect on O_{error} ($\chi^2_{(2,N=72)}$=1.31, p=.52). However, unlike Wobbrock *et al.*'s study, we found no significant effect of W on O_{error} ($\chi^2_{(5,N=72)}$=7.31, p=.20). There was also no significant effect of ID on O_{error} ($\chi^2_{(12,N=72)}$=8.15, p=.77) due to ID's combination with, and the influence of, $MT_{\%}$. Results for predicted error rates agree.

DISCUSSION

On the whole, our error model predicts well in 2-D using either univariate (SD_x) or bivariate ($SD_{x,y}$) endpoint deviation. Despite SD_x's slightly higher R^2, the two were not significantly different in their deviation from observed error rates. In general, error rate predictions using $SD_{x,y}$ were higher than those using SD_x, which makes sense given the extra movement dimension taken into account.

Despite W not causing a significant difference in observed error rates (p=.20), there was still "banding" evident. (A difference here from Wobbrock *et al.* [13] is that they used three levels of W whereas we used six.) Figure 2 shows that error rates for the smallest *W* (8 pixels) stood out from the rest. Figure 3a shows how MT_e and predicted error rates were linear in *ID*, just as Fitts' law requires. But Figure 3b shows that this linear behavior did not emerge when points were grouped by *W*. Indeed, prior work [11] has shown that error rates decrease with increases in *W even when ID is maintained*, suggesting a *W*-specific effect.

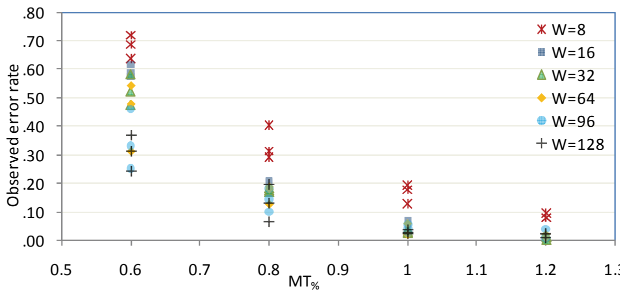

Figure 2. Error rates by $MT_{\%}$ coded by target size *W*. Levels are mixed, but some banding is clear for *W* = 8 (see Figure 7b in [13]).

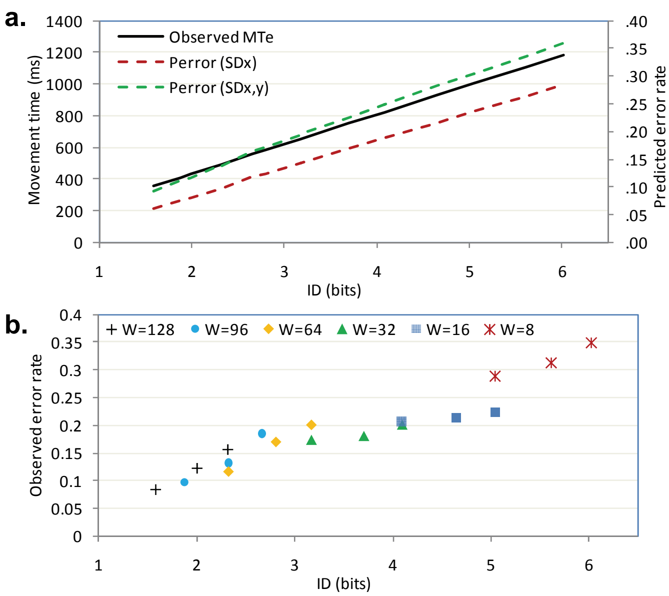

Figure 3. (a) Observed movement times are linear with index of difficulty as predicted by Fitts' law. Predicted error rates follow suit (see Figures 8a-b in [13]). **(b)** Observed error rates show some grouping by *W*, not falling on a line (see Figure 8c in [13]).

CONCLUSION AND FUTURE WORK

In this paper, we have shown that the pointing error model serves well to predict error rates from target distance (*A*) and size (*W*) and actual movement time (MT_e). Both univariate (SD_x) and bivariate ($SD_{x,y}$) endpoint deviations worked well, predicting error rates with similar differences from observed error rates. Although neither *A* nor *W* had a significant effect on observed rates, *W* seems to have a greater effect on error rates than *A*. Future work should investigate *W* further, teasing out any disproportionate effect on errors and adjusting the pointing error model accordingly. Future work should also augment the model for noncircular and irregularly-shaped targets. Performance should be predicted for a game, *e.g.*, *Whack-a-Mole*, where movement is governed by game conditions, not by Fitts' law.

ACKNOWLEDGEMENTS

The authors thank Edward Cutrell and I. Scott MacKenzie. This work was supported in part by the National Science Foundation under grants IIS-0811063 and IIS-0952786.

REFERENCES

1. Crossman, E.R.F.W. (1957). The speed and accuracy of simple hand movements. In *The Nature and Acquisition of Industrial Skills*, Crossman and Seymour (eds.). Report to the Joint Committee on Individual Efficiency in Industry.
2. 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. *Proc. CHI '99*. New York: ACM Press, 215-222.
3. Fitts, P.M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *J. Experimental Psychology 47* (6), 381-391.
4. International Organization for Standardization. (2002). Ergonomic requirements for office work with visual display terminals (VDTs)—Requirements for non-keyboard input devices. *Ref. No. ISO 9241-9:2000(E)*.
5. Isokoski, P. (2002). Speed-accuracy measures in a population of six mice. *Proc. APCHI '02*. Beijing, China: Science Press, 765-777.
6. MacKenzie, I.S. (1992). Fitts' law as a research and design tool in human-computer interaction. *Human-Computer Interaction 7* (1), 91-139.
7. MacKenzie, I.S. and Isokoski, P. (2008). Fitts' throughput and the speed-accuracy tradeoff. *Proc. CHI '08*. New York: ACM Press, 1633-1636.
8. Murata, A. (1999). Extending effective target width in Fitts' law to a two-dimensional pointing task. *Int'l J. Human-Computer Interaction 11* (2), 137-152.
9. 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.
10. 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. Human-Computer Studies 61* (6), 751-789.
11. Wallace, S.A. and Newell, K.M. (1983). Visual control of discrete aiming movements. *Quarterly J. Experimental Psychology 35A* (2), 311-321.
12. Welford, A.T. (1968). *Fundamentals of Skill*. London, England: Methuen.
13. Wobbrock, J.O., Cutrell, E., Harada, S. and MacKenzie, I.S. (2008). An error model for pointing based on Fitts' law. *Proc. CHI '08*. New York: ACM Press, 1613-1622.
14. 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. *Proc. CHI '11*. New York: ACM Press. To appear.