

Physical Versus Virtual Pointing

Evan D. Graham
 School of Kinesiology
 Simon Fraser University
 Burnaby, BC
 Canada V5A 1S6
 (604) 291-5794
 evan_graham@sfu.ca

Christine L. MacKenzie
 School of Kinesiology
 Simon Fraser University
 Burnaby, BC
 Canada V5A 1S6
 (604) 291-3004
 christine_mackenzie@sfu.ca

ABSTRACT

An experiment was conducted to investigate differences in performance between virtual pointing, where a 2-D computer image representing the hand and targets was superimposed on the workspace, and physical pointing with vision of the hand and targets painted on the work surface. A detailed examination of movement kinematics revealed no differences in the initial phase of the movement, but that the final phase of homing in on smaller targets was more difficult in the virtual condition. These differences are summarised by a two-part model of movement time which also captures the effects of scaling distances to, and sizes of, targets. The implications of this model for design, analysis, and classification of pointing devices and positioning tasks are discussed.

Keywords

Analysis methods, Fitts' law, human performance modelling, input devices, pointing, virtual environments.

INTRODUCTION

Pointing to a location on a graphics display is an elemental gesture in many forms of human-computer interaction (HCI). Pointing movements have been studied in an attempt to understand perceptual-motor processes when we interact with real objects in the physical world. Our interest is in relating these theories and models from motor control to human performance in more abstract environments, where objects and actions represented on a graphics display are mediated by pointing devices. In particular, we wonder how limitations of current 2-D and 3-D virtual environments affect planning and control of natural movements like aiming, pointing, reaching,

grasping, and manipulating objects; and how detailed analyses of movement kinematics can be used to reveal systematic effects of these constraints on human performance, in the HCI context.

Woodworth [12] first proposed that human pointing movements can be understood in terms of two movement phases: an initial planned impulse which covers most of the distance, followed by a second phase of deceleration to the target under current control. According to Fitts [3], total movement time involves a tradeoff between speed (distance covering) and accuracy (deceleration to the target). One version of Fitts' law states that on average, total movement time (MT) is a linear function of the index of difficulty (ID) of pointing to a target, where ID, in units of bits, is defined as $\log_2(2A/W)$, A being the distance to the target, and W the width of the target:

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

This relation has been demonstrated in a number of studies of discrete and repetitive tapping and pointing with physical targets, using an tool such as a stylus [4] [2] [11], and also confirmed with computer systems and pointing devices ([1],[10], for a review see [8]) under a wide range of conditions.

In more recent studies of movement kinematics, Soechting [6], and Langolf et al. [9] examined velocity profiles along the path of a pointing movement, showing that the shape of the profiles was skewed for smaller, more difficult targets. MacKenzie et al. [7] showed how this effect was systematic for three-dimensional pointing movements, and suggested that a base trajectory representation of the velocity profile exists for a given size (W) of physical target, and is scaled according to the amplitude (A) of movement. They also proposed that for different combinations of A and W which give rise to the same ID, the shape and scaling effects cancel one another to give equal movement times as predicted by Fitts' law, even though the shape of the velocity profiles vary dramatically

Permission to make digital/hard copies of all or part of this material for personal or classroom use is granted without fee provided that the copies are not made or distributed for profit or commercial advantage, the copyright notice, the title of the publication and its date appear, and notice is given that copyright is by permission of the ACM, Inc. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires specific permission and/or fee.

CHI 96 Vancouver, BC Canada
 © 1996 ACM 0-89791-777-4/96/04..\$3.50

as A and W are scaled in proportion. In a previous study [5] where subjects used 3-D pointing movements to control a cursor on a 2-D computer display with varying control-display gain settings, we demonstrated these separable effects of A and W on the velocity profiles of both the hand and cursor. We found that:

- a) the timing and magnitude of the initial velocity peak were a function of the distance to be covered by the hand, regardless of the scale of the display and target width; and
- b) the proportion of time spent after the first deceleration peak was a function of the target width as an accuracy constraint for the hand, regardless of the size of its graphical representation or distance covered on the display.

In our study, however, we also observed unequal scaling of these amplitude and width effects. The overall size of the velocity profile was more sensitive to changes in amplitude than its change of shape for differing target widths. As a result, movement times were significantly longer to more distant targets with the same index of difficulty. We used a two-part model from Welford [11, p. 157] to summarise these results:

$$MT = a + b_1 \log_2 A - b_2 \log_2 W \quad (2)$$

In Welford's original formulation of the model, the units for b_1 and b_2 were bits per second. For linear regression on our data, we used this simplified form which distinguishes two ways of altering the index of difficulty for a pointing task: changing A or changing W. The first coefficient b_1 , related to distance-covering, gives the sensitivity of MT to changes in A. The second coefficient b_2 , related to deceleration to the target, gives the sensitivity to changes in W. In the case where b_1 and b_2 are approximately equal, Fitts' law would give a good account of the data. In our study of 3-D pointing to control a cursor on a 2-D display, we found that:

$$MT = 129 + 153 \log_2 A - 83 \log_2 W \quad (R^2 = .97) \quad (3)$$

with units of milliseconds and millimetres. We wondered if the large difference between coefficients might be due to a particular aspect of the setup used in the experiment which made the task more difficult than the corresponding task of pointing to a physical target with full vision of the hand. In our study, the subject made pointing movements with the hand on a table top in front and to the right of the body while viewing a display placed directly in front, tilted slightly away from vertical. We refer to this configuration as the "standard HCI setup", typically used with systems employing a mouse or tablet as a pointing device. We felt the main differences between our task and physical pointing were:

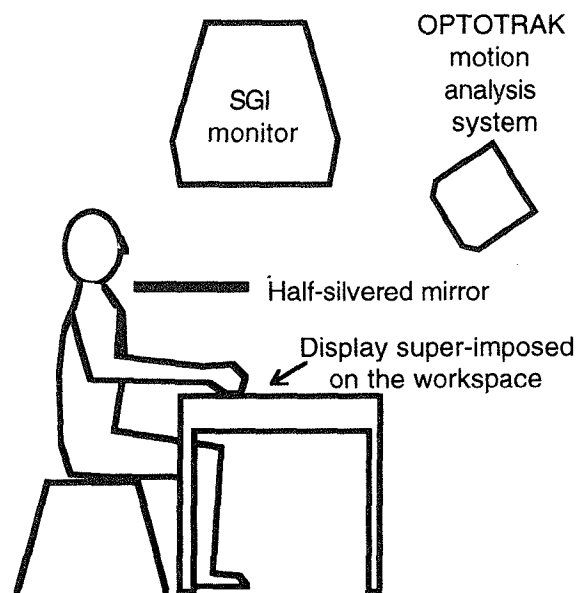


Figure 1. Layout of the experimental setup. Subjects made pointing movements while viewing either a graphics image of targets and a pointer superimposed on the workspace, or a physical pointer mounted on their index finger and targets painted on the table surface. Kinematic data for hand position was captured by an OPTOTRAK motion analysis system.

- 1) In the HCI configuration the display space is physically separate from working space for the hand. Movement planning requires learning a sensorimotor transformation to map the direction and extent of hand movements to corresponding cursor movements on the display.
- 2) The 2-D display presents no visual information about height of the hand above the table top, even though this dimension is important in control of the hand to achieve contact with the target location.

For the present study we eliminated the first difference (1 above) by superimposing a graphics image directly on the working space for the hand, and contrasted virtual (using a 2-D display) and physical (using full vision of the hand and physical targets) versions of the pointing task.

METHOD

Subjects

Six volunteers from the local community were paid to participate in the experiment. All had normal or corrected vision, had experience with computer pointing devices, and preferred to use the right hand in this context.

Apparatus

Subjects were comfortably seated in front of a table, viewing the work surface through a half-silvered mirror as shown in Figure 1. An OPTOTRAK motion analysis system recorded the three-dimensional position of infrared markers (IREDS) placed on the index finger, hand and arm, and stored the results in data files for further analysis. Coordinates from the OPTOTRAK were also used in real time by a Silicon Graphics Indigo Extreme workstation to update a 19" RGB monitor, double buffered at 60 frames per second. Sensing lag due to the OPTOTRAK was approximately 5 ms., and both the pointer and targets were represented as flat polygons on the display. As a result, the total system lag from hand (through OPTOTRAK, workstation, and display) to the eye was verified to be no greater than 25 ms, or 1 1/2 frames at 60 Hz. Subjects performed two versions of the pointing task:

1) Virtual pointing, where the half-silvered mirror was blocked so that only the graphics display image was visible. Targets were represented as white circles on a black background, while the two IRED markers on the index finger were used to superimpose an image of a red pencil-shaped "virtual finger" on the actual position of the subjects' finger over the work surface.

2) Physical pointing, where the graphics display was turned off and the mirror was unblocked, so that subjects could see through the mirror to the workspace below, illuminated by a task light. White circular targets were painted on a black background to correspond to the images generated by the graphics display for the virtual task. The subjects wore a fluorescent red pencil-shaped pointer, having the same dimensions as the virtual finger above, on their index finger.

Procedure

For each trial, a subject held the tip of the red pointer on a start mark. On instruction from the experimenter, the subject lifted the hand and quickly pointed to a spot on the table surface so that the tip of the red pointer ended anywhere inside the target circle. Movement speed was emphasised, with small corrections allowed if the target was missed. Four target sizes ($W = 6, 12, 24, 48$ mm) were combined, with four movement distances ($A = 37.5, 75, 150, 300$ mm) for a total of sixteen A and W combinations.

Data for each subject were collected during a single session lasting about 1 1/2 hours. After a practice session with one of the display configurations, subjects performed twelve movements in each A and W combination, with a short break between each set of twelve trials while the target was changed. The order of presentation of A and W combinations was randomised, and the order of presentation

Table 1. Summary of main effects for display.

Display	Virtual	Physical
MT (ms)	610	538 **
PA (mm/s/s)	7688	8514
TPA (ms)	100	97.8
PV (mm/s/s)	721	751
TPV (ms)	192	186
%TAPD	56.1	53.7

** $p < .01$

of the virtual or physical task was counterbalanced across subjects.

Data Analysis

Three-dimensional position data for each trial were rotated and translated to a convenient coordinate system where: the x-axis represents the principal direction of movement, away from the subject; the y-axis points to the left of the subject, and the z-axis points up. Data were then smoothed by a low-pass second order bidirectional Butterworth digital filter with a cutoff of 8 Hz. The three-dimensional trajectories of the fingertip were numerically differentiated to produce a velocity profile, and the resultant (speed of the movement along its path) further differentiated to produce an acceleration profile.

A series of computer programs were used on the resultant of the velocity profile to calculate movement time (MT), and the timing (TPV) and magnitude (PV) of the first velocity peak. Another program used the acceleration profile to calculate the timing (TPA) and magnitude (PA) of the initial acceleration peak, and also the per cent of movement time after the first deceleration peak (%TAPD). The variability of the end point was computed in both the principal direction of movement (VEx) and orthogonal to the principal direction of movement (VEy). VE measures were used to determine [8] the effective target widths (W_e), where $W_e = 4.133 VE$.

The within subject means of the last ten good trials in each A and W combination were analysed using ANOVA with repeated measures (BMDP 8V and 2V) on a subjects (6) by display (2) by movement amplitude (4) by target width (4) design. We also performed separate multiple regressions (BMDP 2R) on means for MT as a function of A, W, and W_e for the model of equation 2 for both virtual and physical pointing.

Table 2. Summary of main effects for movement amplitude.

A (mm)	37.5	75	150	300	
MT (ms)	420	510	613	753	***
PA (mm/s/s)	5148	6114	9187	11957	***
TPA (ms)	89.1	99.3	104	104	***
PV (mm/s/s)	324	471	837	1312	***
TPV (ms)	143	177	201	236	***
%TAPD	57.1	53.6	54.3	54.6	

*** p < .001

Table 3. Summary of main effects for target width.

W (mm)	3	6	12	24	
MT (ms)	657	607	536	496	***
PA (mm/s/s)	7487	8041	8672	8806	**
TPA (ms)	97.1	99.5	98.7	101	
PV (mm/s/s)	719	747	712	766	
TPV (ms)	190	190	190	187	
%TAPD	61.7	57.0	51.9	49.1	***

** p < .01, *** p < .001

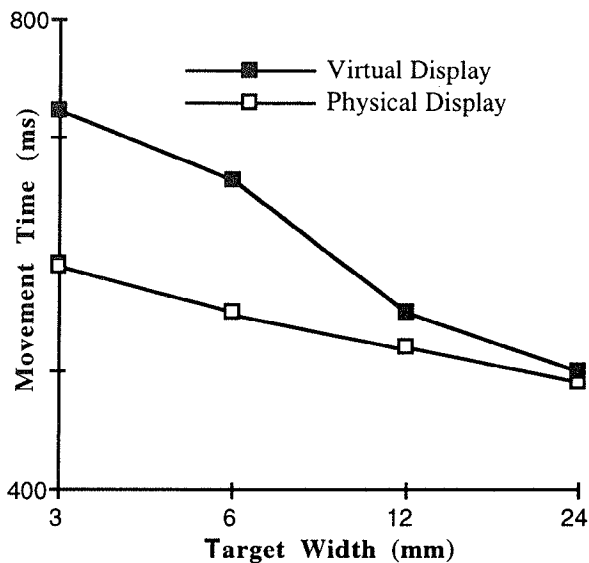


Figure 2. Mean movement times as a function of target width for virtual and physical displays, showing a display by target width interaction. Target width effects on MT are more pronounced for the virtual than the physical display.

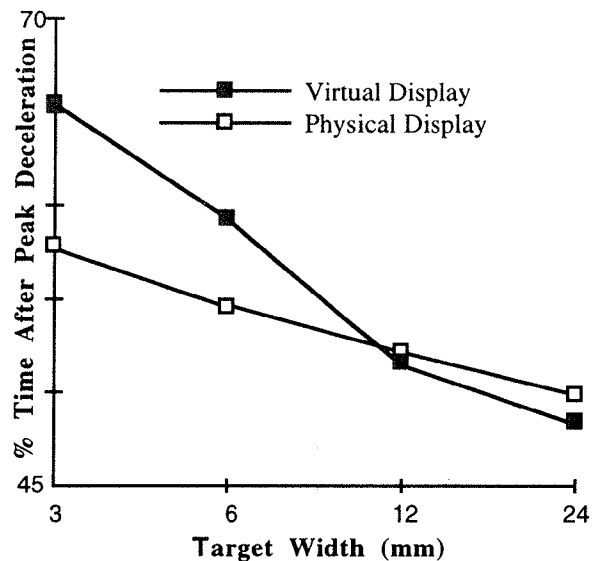


Figure 3. Mean proportion of time spent in deceleration to targets of different widths for virtual and physical display, showing a display by target width interaction. Target width effects on %TAPD are more pronounced for the virtual than the physical display.

RESULTS

Overall differences in 3-D kinematic measures for the hand due to the display are shown in Table 1. Movement times were about 80 ms slower ($F_{1,5}=25.35, p < .01$) with the virtual display than for physical pointing. Differences in MT were evident in a display by target width interaction ($F_{3,15} = 22.9, p < .01$). Shown in Figure 2, the increase in movement times as target size decreased was more evident for the virtual display compared to the physical display condition.

There were no significant differences between the physical and virtual display condition for the initial movement

phase, up to the time of peak velocity. The slightly longer proportion of movement time (%TAPD) for the virtual display (56.1 versus 53.7 per cent) was not statistically significant, but revealed itself in a display by width interaction ($F_{3,15}=4.97, p < .02$) as shown in Figure 3. For smaller targets, the %TAPD increased more with the virtual display than with the physical display. In contrast, the %TAPD was similar across displays for the largest targets. This interaction parallels the interaction for MT, and demonstrates that the longer movement times for virtual compared to physical pointing are mainly due to extra time in deceleration to smaller targets when the virtual display was used.

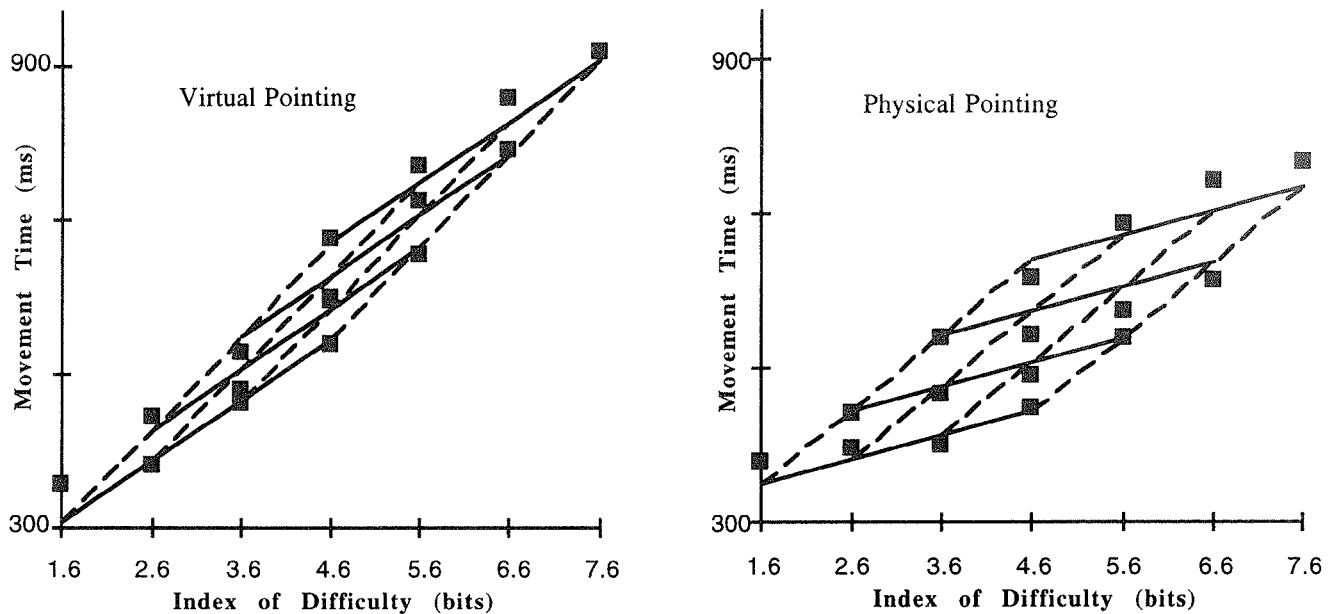


Figure 4. On the left, the two-part model of equation 4 is superimposed on actual data points for virtual pointing. Similarly on the right, the two-part model of equation 5 is superimposed on actual data points for physical pointing. Movement times (black squares) are plotted as a function of index of difficulty (ID). Dashed lines show predicted MT when target width is held constant, and ID is changed by varying A. Solid lines show predicted MT when amplitude is held constant, and ID is changed by varying W.

The main effects for movement amplitude are shown in Table 2. As predicted by Fitts' law, movement time increased for longer distances. Differences in the velocity profile due to movement amplitude are evident up to the time of the first velocity peak. In particular, the magnitudes of the initial acceleration and velocity peaks (PA and PV) increase with longer distances, and the timing of these peaks is also delayed for longer distances in a systematic fashion. All these effects were highly significant ($p < .001$). Although there were small differences in %TAPD for different movement amplitudes, this effect did not approach statistical significance.

Main effects for target width are shown in Table 3. As expected, movement time decreased for larger targets. Target size showed little effect on the first movement phase, although there was a slight increase in the magnitude of the first acceleration peak ($F_{3,15} = 5.59$, $p < .01$) for larger targets. The proportion of movement time spent in deceleration increased systematically from 49.1 to 61.7 per cent as target size decreased ($F_{3,15} = 22.14$, $p < .001$) from the large to small. We captured the effects of A and W on movement time using multiple linear regression to fit the model of equation 2 to means of MT for each A and W combination (units are millimetres and milliseconds). For virtual pointing:

$$MT = 22 + 123 \log_2 A - 79 \log_2 W \quad (R^2 = .99) \quad (4)$$

and for physical pointing:

$$MT = -24 + 98 \log_2 A - 32 \log_2 W \quad (R^2 = .96) \quad (5)$$

Note that in both conditions MT is more sensitive to changes in A (virtual 123 ms/bit, physical 98 ms/bit) than in W (virtual 79 ms/bit, physical 32 ms/bit), and this asymmetry is even more marked for the physical display than for virtual pointing. One explanation for this phenomenon is that subjects were not taking advantage of the relaxed accuracy constraint to improve their speed in reaching larger targets, so that effective target width was not varying as much as the actual width presented during trials. Since our analysis of variable error (VE) revealed the greatest deviations along the principal axis of movement (x), we used VEx to compute an effective target width (W_e) for each amplitude and width mean. Modelling MT with W_e rather than W revealed the same pattern of coefficients as equations 4 and 5 (units are mm and ms). For virtual pointing:

$$MT = 30 + 132 \log_2 A - 89 \log_2 W_e \quad (R^2 = .93) \quad (6)$$

and for physical pointing:

$$MT = -28 + 101 \log_2 A - 41 \log_2 W_e \quad (R^2 = .95) \quad (7)$$

This indicates two things: first, that the subjects were

taking advantage of larger targets to increase the movement speed, and second, that equations 4 and 5 effectively describe subjects' performance in the task.

To visualise the two-part model of equation 4, in Figure 4 on the left we have plotted actual and predicted movement times against ID for virtual pointing. Similarly, the model of equation 5 is depicted in Figure 4 on the right for physical pointing. As explained in the figure caption:

a) Movement times are plotted as black squares. Differences for various A and W combinations giving rise to the same ID are evident by the vertical spread of the data points.

b) The dashed lines represent predicted movement times when target width is held constant, and ID is changed by varying A.

c) The solid lines represent predicted movement times when amplitude is held constant, and ID is changed by varying W.

Looking at Figure 4, two features are evident. First, changing A (dashed lines) gives rise to a steeper slope than changing W (solid lines) for both display conditions. Second, the difference in slopes for A (dashed lines) and W (solid lines) is more marked for physical pointing than for virtual pointing.

DISCUSSION

Theory

The results suggest that movement planning is similar for both virtual and physical pointing in our experiment, evidenced by the lack of differences in kinematic features of the initial phase of the movement. The difference between the virtual and physical display is apparent only in the second movement phase, where visual control of deceleration to the smaller targets in the virtual task took more time than in the physical task. This interpretation is also supported by contrasting the models of equations 4 and 5: The coefficients (b_1) for the A term are of similar magnitude, but there is a marked difference in the coefficients for the W term (b_2), making MT much more sensitive to the increased accuracy constraint of a smaller target in virtual pointing.

A key feature of our virtual display is that it is flat — it presents no visual information to show the height of the finger above the table top, even though this dimension is used as part of the deceleration strategy to contact the table surface at the target location. This lack of the third dimension could account for the difficulty with visual control during final positioning in virtual pointing, and

suggests further study using a more realistic 3-D virtual environment to clarify this issue.

To our surprise, we found that even with physical pointing changes in the size and shape of the velocity profile did not cancel when A and W were increased in proportion, as would be predicted by Fitts' Law. The spread of movement times for similar IDs in physical pointing was even greater than either virtual pointing, or our previous results using the standard HCI configuration. Thus, the fact that movement times increase as A and W are scaled up in size seems to be a feature of the pointing task per se, not a result of limitations in motor control processes associated with using a computer display.

Although the two-part model due to Welford (equation 2) is difficult to interpret in terms of an information-theoretic account of movement planning and control processes, it does capture the large and significant differences in movement time due to scaling. We suggest that the two-part models captured in equations 4 and 5 promise some utility from an engineering perspective. First, as a predictive model of performance, it takes into account the size or scale of hand movements appropriate for a particular interface. Second, as a diagnostic, it can serve to identify the type of constraint leading to speed-accuracy tradeoffs in a variety of positioning tasks.

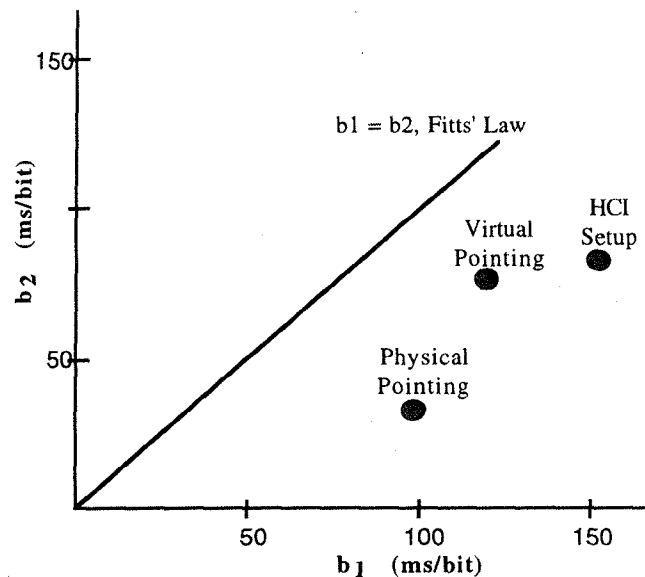


Figure 5. Parameter space for the model of equation 2, showing the data points for virtual and physical pointing, and also for pointing using the HCI setup in a previous study.

Application

In a graphical interface design, there is often some leeway in the choice of size both for a visual representation, and for the hand movements to operate an interface. For example, a set of buttons can be sized and spaced in various ways on the screen. There are obvious limitations on the lower limit of size, such as visual acuity and pointing device and display resolution, as well as on the upper limit, such as the footprint for working with a pointing device, but within these bounds the designer has to make an informed choice.

Figure 5 represents a parameter space for the coefficients of the two-part model of equation 2. In this figure we have plotted the data points for virtual and physical pointing from equations 4 and 5. In addition, we have added the point for our previous study [5], showing where equation 3, representing the standard HCI setup, falls in the parameter space. If the data from our studies were well characterised by Fitts' law, they would fall on the line $b_1 = b_2$. In our case, all points fall below this line by different amounts. In this region, below the Fitts' law line, movement time increases as the movement distance (and target width) for the hand are scaled up in size. Tasks associated with points in this region of the parameter space ($b_1 > b_2$) will be performed faster when hand movements are scaled to a smaller size - smaller is better. A glance at the points on the graph of Figure 5 tells us that our tasks will be performed faster if the movement distance and target size are reduced to the smallest practical value. In contrast, if we were to find a two-part model for a task which fell above the Fitts' law line ($b_1 < b_2$), this task would benefit from being scaled up - bigger is better.

The parameter space of Figure 5 may also be useful to classify pointing devices and display systems. Approximate values for the two-part model coefficients for a particular configuration can be readily identified by testing a small number of representative users with a large and small target, and a large and small movement amplitude, to form a total of four A and W combinations. We suggest that data points that fall higher on the graph (b_2 is large) indicate a task or device which involves more difficulty in final positioning on the target. A data point which falls more to the right (b_1 is large) indicates difficulty in planning and control of the distance covering phase of the movement with that system.

We plan to follow up these results in two ways. First, the study described here will be repeated using a 3-D stereographic display superimposed on the workspace, in order to see if the differences between the virtual and physical task we reported are due to lack of visual information about the height of the hand above the workspace in our 2-D virtual environment. Second, we are

compiling data from some of our other studies, as well as revisiting data reported both in the HCI and motor control literature on discrete and repetitive pointing, in order to model and plot additional points in the parameter space shown in Figure 5. We hope to develop this approach into a useful technique for characterising and quantifying human performance for different combinations of positioning task and input device.

ACKNOWLEDGEMENTS

The authors thank Anita Trip for her assistance in data collection. This research was supported from research and strategic grants from the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

1. Boritz, J., Booth, K.S., and Cowan, W.B. (1991). Fitts' law studies of directional mouse movement. In *proc. of Graphics Interface '91*. Toronto: Canadian Information Processing Society, 216-223.
2. Crossman, E.R.F.W., and Goodeve, P.J. (1983). Feedback and control of hand movement and Fitts' law. *Quarterly Journal of Experimental Psychology*, 35A, 251-278.
3. Fitts, P.M. (1954). The information capacity of the human motor system in controlling the amplitude of movement. *J. Exp. Psych.*, 47(6), 381-391.
4. Fitts, P.M., and Peterson, J.P. (1964). Information capacity of discrete motor responses. *J. Exp. Psych.* 67(2), 103-112.
5. Graham, E.D. and MacKenzie, C.L. (1995). Pointing on a computer display. *CHI '95 Conference Companion*, Denver, Colorado, May, pp. 314-315.
6. Langolf, G.D., Chaffin, D.B., and Foulke, J.A. (1976). An investigation of Fitts' law using a wide range of movement amplitudes. *J. of Motor Behaviour*, 8, 113-128.
7. MacKenzie, C.L., Marteniuk, R.G., Dugas, C., Liske, D., and Eickmeier, B. (1987). Three-dimensional movement trajectories in Fitts' task: implications for control. *Q. J. Exp. Psych.*, 39A, 629-647.
8. MacKenzie, S. (1992). Fitts' law as a research and design tool in human-computer interaction. *Human-Computer Interaction*, 7, 91-139..
9. Soechting, J.F. (1984). Effect of target size on spatial and temporal characteristics of a pointing movement in man. *Experimental Brain Research*, 54, 212-132.

10. Walker, N., Meyer, D.E., Smelcer, J.B. (1993). Spatial and temporal characteristics of rapid cursor-positioning movements with electromechanical mice in human-computer interaction. Human Factors, 35(3), 431-458.
11. Welford, A.T. (1968). Fundamentals of Skill. London: Methuen.
12. Woodworth, R. (1899). The accuracy of voluntary movement, Psych. Review Monograph Supplement, 3.