Quantitative Analysis of Scrolling Techniques

Ken Hinckley, Edward Cutrell, Steve Bathiche, and Tim Muss
Microsoft Research, One Microsoft Way, Redmond, WA 98052
{kenh, cutrell, stevieb, timmuss}@microsoft.com

ABSTRACT

We propose a formal experimental paradigm designed to help evaluate scrolling interaction techniques. Such a method is needed by interaction designers to quantify scrolling performance, thereby providing a tool to evaluate and improve upon new techniques. We systematically vary the scrolling *distance* as well as the required *tolerance* of scrolling. Distance and tolerance are the parameters of Fitts' Law, which traditionally has been applied to the evaluation of pointing devices in tasks involving rapid, aimed movement to visible targets. Scrolling involves acquisition of targets well beyond the edges of the screen, yet Fitts' Law models our experimental data very well.

We apply our paradigm to the IBM ScrollPoint and the IntelliMouse Wheel. Our experimental approach reveals a crossover effect in performance versus distance, with the Wheel performing best at short distances but the ScrollPoint performing best at long distances. We also demonstrate that the performance of the Wheel can be significantly improved using an acceleration algorithm. These results show that our approach yields a practical and rigorous method for the evaluation of scrolling techniques.

Keywords

Document navigation, scrolling, Fitts' Law, C:D gain

INTRODUCTION

Scrolling through a document is a common task performed by millions of computer users every day. Mice with dedicated controls for scrolling, including wheels and isometric joysticks, are now nearly ubiquitous. But despite the market penetration of such devices, we believe that scrolling performance can be significantly enhanced. New techniques continue to appear [10] and the increasing importance of mobile devices has inspired new approaches, such as using tilt to scroll [9]. In short, we are perhaps more in need of a sound method to evaluate scrolling techniques than ever before, but the literature has few empirical studies to offer. This is especially true when compared to the literature for pointing devices, where Fitts' Law is a well known quantitative method for evaluating, optimizing, and studying properties of pointing devices (e.g. [4][6][13]).

The research and design communities need related techniques for the evaluation, analysis, and refinement of scrolling methods. The few empirical studies that do exist (e.g. [3][18]) suffer a common limitation: to our knowledge, there has never been an analysis of scrolling techniques that systematically varies the scrolling distance and the required precision of scrolling. Our results suggest that there can be performance differences between scrolling techniques as a function of distance or tolerance. Hence failure to control for these influential factors yields an incomplete view of a scrolling technique's overall performance, which in turn has the potential to yield erroneous conclusions. To address this significant limitation in the literature, we contribute an experimental paradigm that uses Fitts' Law to study scrolling.

Currently submitted to
CHI 2002

To illustrate our approach and its significance, we analyze several scrolling techniques. For example, previous results suggest that the isometric joystick of the IBM ScrollPoint is faster than the scrolling wheel [18], but our approach shows that the scrolling wheel is actually faster when scrolling distances less than about 100 lines (see Fig. 6). Hence each device has quantifiable strengths and weaknesses. We also show that it is possible to improve performance of the wheel using an acceleration algorithm. When acceleration is used at a high resolution of 1 line per wheel notch, it is better than or equal to the ScrollPoint or unmodified wheel up to about 200 lines. At 3 lines/notch, this increases to about 400 lines (the largest distance that we tested).

RELATED WORK

Multi-stream input devices include mice with a wheel or isometric joystick [18] for scrolling. Zhai et al. report data suggesting that the isometric joystick is 29% faster than a wheel, for the task of visually scanning a web page for a randomly placed but prominent hyperlink ("NEXT") in a text document about 12 pages long. In their study, the wheel was set to move 1 line of text per notch, which is more precise but slower than the manufacturer's default of 3 lines per notch. Subjects were also free to either roll the wheel, or use a rate-control mode available on the IntelliMouse, so it is unclear exactly what mixture of the control modes results in the reported 29% difference.

We are not aware of any previous study of document navigation that controls for possible effects of distance or tolerance: only overall means are reported. This is a potential problem, as shown in Fig. 1, which demonstrates how a hypothetical crossover effect by distance might change the results of an experiment comparing two scrolling techniques. If the average distance in the experiment is **A**, then the "dashed" technique is fastest. At distance **B**, there is no difference between the techniques. But for distance **C**, the "solid" technique is fastest!

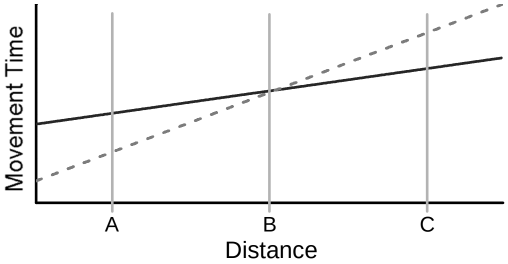


Fig. 1 Influence of uncontrolled crossover effects on a hypothetical experiment.

Fitts' Law has been widely applied to pointing devices, but to our knowledge it has never been applied to scrolling, although it has yielded insight into multi-scale interfaces [8]. When the target is visible, scrolling a short distance (e.g. to precisely place a figure) is analogous to target selection using an area cursor, a task for which Fitts' Law has been shown to apply [12], but there is no direct evidence that shows if Fitts' Law is valid for scrolling.

Regarding the prediction of movement times with traditional pointing devices, MacKenzie [13] states "it is noteworthy of the [Fitts] model in general that correlations above .90 consistently emerge." For our scrolling task, which is certainly not a traditional pointing task, we find that Fitts' Law also models our data with correlations of .90 or more for all of the scrolling devices that we tested. This suggests that Fitts' Law may indeed be relevant to the study of scrolling movements, although further studies are clearly warranted to explore this issue.

Several researchers have argued that the background task of navigating a document should be assigned to the nonpreferred hand, while the preferred hand operates the mouse [3][8]. Buxton and Myers report that a two-handed approach is about 25% faster than scrollbars for novices [3]. We are currently exploring several two-handed document navigation techniques, including those supported by the recent Microsoft Office Keyboard [15], which includes a scrolling wheel on the left side of the keyboard. We expect to report on two-handed techniques in future work, but they are not a focus of this paper.

SCROLLING INTERACTION TECHNIQUES

Our experiment includes the IBM ScrollPoint Pro mouse with isometric joystick (the "ScrollPoint") and the Microsoft IntelliMouse Explorer with scrolling wheel. We chose the ScrollPoint and IntelliMouse wheel as they represent widely available commercial devices which have been included in a previous study [18].

IBM ScrollPoint Pro

The IBM ScrollPoint Pro is a recent commercial version of the prototype implementation described by Zhai et al. [18]. The device scrolls the document at a speed proportional to the force exerted on the isometric joystick. The force sensor used in the ScrollPoint Pro isometric joystick has a lower sensing resolution than the IBM Trackpoint that was used in the original prototype, to allow cost reduction to a level feasible for integration with a mouse. We tested the device using the manufacturer's default settings.

Perhaps the most crucial property of the ScrollPoint is the self-centering nature of the isometric joystick. This allows it to have an unlimited range since holding the stick continuously scrolls the document, but as soon as the stick is released scrolling stops. This is a potential advantage over position-sensing scrolling mechanisms such as the wheel, which have a limited range of movement before the user has to release the device and stroke again. Thus the "unlimited range" should be particularly beneficial for scrolling long distances in a document.

However, rate control devices may be less intuitive than position-control devices such as the wheel, since a higher order transfer function is required to translate the force into a movement [1]. Another problem with rate control devices is that it is difficult to move to a new position, and then quickly flip back to the previous position [1]. This can arise naturally in scrolling movements if the user refers to a nearby section of a document, and then quickly reverses the scrolling to resume editing a paragraph. For short scrolling distances, this can be easily accomplished on a position sensing device such as the wheel by rolling a bit, and then quickly returning the wheel to its previous position.

IntelliMouse Scrolling Wheel

The scrolling wheel on the Microsoft IntelliMouse Explorer is similar to the wheel tested by Zhai et al. [18], although it rolls with less friction. The wheel senses 18 positions per revolution, each felt by a tactile "notch" as one rotates the wheel. We tested the wheel at the manufacturer's *default setting of 3 lines per notch* rather than 1 line/notch as used by Zhai et al. [18]. One line/notch is not the default setting, and it is fatiguing and impractical to scroll hundreds of lines at this setting, so there is little point in testing this. Henceforth we refer to the default 3 lines/notch as the "**Standard Wheel**" configuration.

The IntelliMouse wheel is integrated with a button, which by default triggers a rate scrolling mode that causes scrolling to move increasingly faster as the user moves the mouse further. It is known that rate control mappings for position sensing devices can lead to degraded performance [17]. We have experimented with alternate position-to-position mappings that we believe may improve performance in this mode, but this issue is not a focus of this paper. In fact, we placed a shim under the wheel button to physically disable the button, so that we could study performance of the rolling the wheel itself. This ensures a clean experimental result that is not tainted by uncertainty of which method participants used to scroll.

Wheel Acceleration Algorithm

When rolling the wheel, users tend to exhibit two distinct behaviors. When reading or moving short distances, users move the wheel slowly in a controlled, line-by-line manner. But when covering longer distances, users will often rapidly "flick" the wheel to get there quickly. A probability distribution of wheel speed (measured as the arrival time between individual notches of wheel movement) illustrates these behaviors (Fig. 2, solid curve). The sharp peak on the left represents rapid wheel motions, while the second peak represents slower "reading" motions.

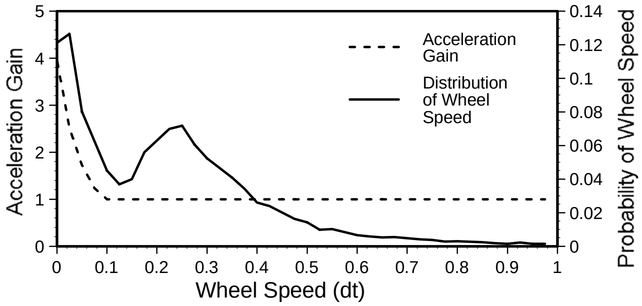


Fig. 2 Bimodal distribution of wheel movements and our acceleration curve plotted on the same time scale.

We generated this graph from the data logs for all users of the Standard Wheel in the experimental task described in this paper. We have observed a similar distribution in data collected from the day-to-day activity of people using the wheel during actual work as well, so this distribution is not an artifact of our experimental task. In fact, one can model the observed distribution as the sum of two Poisson arrival processes (one for fast and one for slow wheel motions) with over 97% of all variance explained (r^2 =.974).

Our acceleration algorithm uses this bimodal distribution to improve performance. We apply a continuous exponential transformation to rapid scrolling movements, but slow scrolling movements are not accelerated (Fig. 2, dotted curve). The cutoff threshold is at Δt=0.1 seconds, which is just to the left of the local minima between the two "peaks" of the probability distribution. This gives the algorithm a slight conservative bias *not* to accelerate a movement which the user may have intended as a slow, controlled motion rather than a rapid flicking motion. The exponential transformation has the form:

$$\Delta y = K_1(1 + K_2\Delta t)^{\alpha} \qquad \text{(Equation 1)}$$

Where Δy is the resulting scroll gain factor, Δt is the time difference between subsequent wheel notches, K_1 and K_2 are constants, and α is the exponent. With the parameters K_1=4, K_2=8, and α = -2.5, the acceleration seems to aid performance but also occurs gradually, allowing the user to visually track a document without disorienting jumps. Note that fractional lines of scrolling may follow from Eq. 1.

As the algorithm's design is based on a naturally occurring behavior pattern, users do not have to learn anything new to benefit from this enhancement to the wheel. Certainly, one of our motivations for the present research was to quantify performance of our algorithm and determine if it was always beneficial or if it might possibly hinder performance in some task conditions (e.g., it might speed up long movements at the expense of short ones).

In this experiment, we evaluate acceleration for the wheel at 3 lines/notch ("**Accel W3**") and at 1 line/notch ("**Accel W1").** We tested Accel W3 to allow a direct comparison to the unaccelerated Standard Wheel, which also uses a 3 line/notch setting. We tested Accel W1 to see if acceleration might allow the device to perform effectively for long distance scrolling while also enabling the user to scroll in smaller increments if desired.

Our acceleration algorithm is specific to the wheel, but our approach of studying patterns in user performance and exploiting those patterns to design improved input mappings is a general strategy that can be applied elsewhere. Also, note that we are not aware of any "acceleration" techniques that might improve performance of the ScrollPoint, as its properties differ from the wheel. It may be possible to devise future improvements, but designing a good transfer function for force-sensing devices is known to be a challenging problem [16].

FITTS' PARADIGM AS A TOOL TO STUDY SCROLLING

It has been stated clearly in the literature that "we must recognize when not to use Fitts' Law. The law is a prediction model for rapid, aimed movement" [14]. Indeed by this description it might seem unlikely that Fitts' Law should be relevant to scrolling, which often involves acquisition of targets that are not yet visible on the screen.

A movement follows Fitts' Law if the equation:

$$MT = a + b \log_2(D/W + 1) \qquad \text{(Equation 2)}$$

satisfactorily models observed behavior. In Fitts' Law, MT is the movement time, D is the distance of the movement, and W is the width of the target that must be selected. The constants a and b are regression coefficients fit to observed data for MT. The term $\log_2(D/W + 1)$ is also known as the Index of Difficulty (ID). In typical rapid, aimed movement studies, it is not unusual to see Fitts' Law model the mean observed movement times with more than 85% of the variance explained by Eq. 2 [13].

However, it is important to note that this all relates to Fitts' Law as a prediction model. Our primary goal is not to predict the movement time of a scrolling action; rather it is to develop a useful evaluation tool that can quantify relative performance differences among scrolling techniques. We assert that Fitts' task paradigm can be used as an *experimental task* whether or not Fitts' Law actually does model scrolling movements. We believe Fitts' task paradigm provides a useful barometer to evaluate, compare, and make design improvements to scrolling interfaces. Our data suggests Eq. 2 can predict the resulting movement times, but the application of Fitts' Law to *predict* scrolling movement times in general is more of a theoretical issue that will need further exploration in future studies.

REPRESENTATIVE TASKS FOR SCROLLING

When selecting a pointing device, experts recommend to "experiment with a diverse set of representative tasks, each [with] its own idiosyncratic demands" [2]. Such representative tasks are well known for pointing devices, but are lacking for scrolling techniques. When we began evaluating different scrolling techniques, we explored a number of experimental tasks, which reflect various user activities involving scrolling. For example, we tried:

- Scrolling while proofreading text for misspellings.
- Searching for a highlighted line in a document.
- Searching for a highlighted target word in a document, in the presence of highlighted distracter words.

In pilot studies, these tasks yielded useful comments about the techniques, even though quantitative data was not always informative. Of the tasks we explored, we felt that the Fitts' reciprocal task paradigm described below was the most sensitive to differences between techniques, which is why we focus on it here. However, other representative tasks should be useful as part of an overall evaluation of a scrolling solution, even though our emphasis in this paper is on the rigorous quantitative results that can be obtained with our Fitts' Law experimental and task paradigm.

THE EXPERIMENT

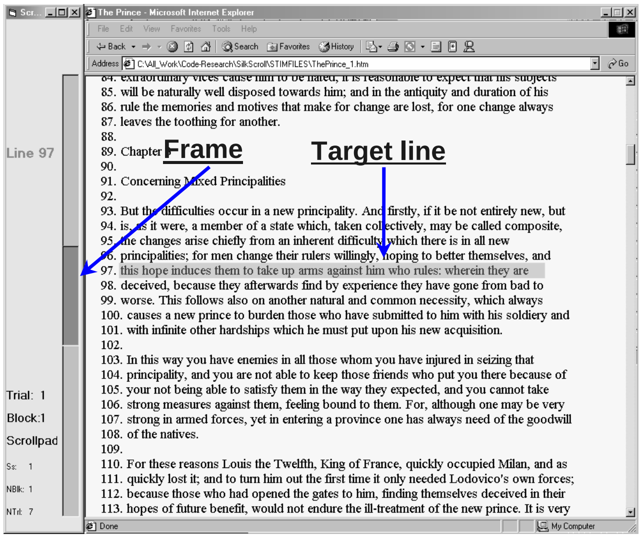
Reciprocal Framing Task for Scrolling

We used a variant of Fitts' reciprocal tapping task adapted to suit vertical scrolling movements. In Fitts' reciprocal tapping task, subjects tap (or point at) two targets as quickly and accurately as possible, moving back and forth between them. In our task, subjects scrolled down, then scrolled up, moving back and forth between two numbered lines in a document using one of the devices (other methods of scrolling, e.g., the scrollbar, were disallowed). A similar task should also be suitable for horizontal scrolling, although we have not yet tried this. We colored the initial target line red and the second target line blue.

The "frame" at the left of the task window (see Fig. 3) specified the tolerance for the target selection, and was colored to match the current target line (red or blue). For each target, subjects scrolled until the target entered the range of the screen identified by the frame. The frame was always centered on the screen (15 lines from the top of the page), and ranged from a tolerance of W=6 to W=30 lines. The visible portion of the document measured exactly 30 lines, so in the limit of W=30, a target was considered to be "in the frame" as soon as it was visible on the screen.

Once the target line was fully within the frame, the subject hit the *selection key* (Caps Lock) with the left hand. This selection key let the computer know when the user judged the scrolling to be complete. We also tried using a dwell time for selection, but this was unsatisfactory as it introduced an arbitrary wait time that might have interfered with task performance; it also precluded measurement of error rates. If the target line successfully fell within the frame when the user struck the selection key, there was a short "happy" sound. If not, the user heard a "bad" sound, but we instructed subjects to always continue with the next target (rather than trying to repair the error).

Fig. 3 illustrates the task. Here the first target was line 97, and the frame width was 6 lines. Subjects scrolled until line 97 fell within the frame, and then tapped the selection key[1]. They then immediately scrolled to and selected the paired target line (e.g., line 121, yielding D=24 lines), after which they returned again to line 97, and so on. Subjects completed at least 10 individual target acquisitions (*phases*) for each trial.



C:\All_Work\Code-Research\SilkScroll\STIMFILES\ThePrince_1.htm

Line 97

Trial: 1
Block:1
Scrollpad
Ss: 1
NBlk: 1
NTrl: 7

84. extraordinary vices cause him to be hated, it is reasonable to expect that his subjects
85. will be naturally well disposed towards him; and in the antiquity and duration of his
86. rule the memories and motives that make for change are lost, for one change always
87. leaves the toothing for another.
88.
89. Chapter
90.
91. Concerning Mixed Principalities
92.
93. But the difficulties occur in a new principality. And firstly, if it be not entirely new, but
94. is, as it were, a member of a state which, taken collectively, may be called composite,
95. the changes arise chiefly from an inherent difficulty which there is in all new
96. principalities; for men change their rulers willingly, hoping to better themselves, and
97. this hope induces them to take up arms against him who rules: wherein they are
98. deceived, because they afterwards find by experience they have gone from bad to
99. worse. This follows also on another natural and common necessity, which always
100. causes a new prince to burden those who have submitted to him with his soldiery and
101. with infinite other hardships which he must put upon his new acquisition.
102.
103. In this way you have enemies in all those whom you have injured in seizing that
104. principality, and you are not able to keep those friends who put you there because of
105. your not being able to satisfy them in the way they expected, and you cannot take
106. strong measures against them, feeling bound to them. For, although one may be very
107. strong in armed forces, yet in entering a province one has always need of the goodwill
108. of the natives.
109.
110. For these reasons Louis the Twelfth, King of France, quickly occupied Milan, and as
111. quickly lost it; and to turn him out the first time it only needed Lodovico's own forces;
112. because those who had opened the gates to him, finding themselves deceived in their
113. hopes of future benefit, would not endure the ill-treatment of the new prince. It is very

Fig. 3 Screen capture of reciprocal framing task, showing target line (97) within the frame, W=6 lines (126 pixels).

Each line of text in our document was 21 pixels (0.59 cm) high. The visible portion of the document was 17.7 cm high by 23.8 cm wide. For the document content, we used four selections from public domain books, cropped at 600 lines long. Note that Internet Explorer (IE) moves 132 pixels per notch for a wheel set at the default of 3 lines/notch. Hence each notch of the Standard Wheel corresponded to 6.29 lines in our document. We also disabled IE's "smooth scrolling" feature, which animates scrolling movements[2].

Following the example of Buxton and Myers [3], we used numbered lines to simulate scrolling in a familiar document. This allowed subjects to find the target lines without excessive searching, yielding a very sensitive measurement of the manual costs of scrolling. We plan to explore the cognitive and attentional costs of scrolling techniques as well, but it is essential to first understand the differences between scrolling input mechanisms in terms of the manual costs, without introducing uncontrolled factors that might occur in a totally unfamiliar document.

Subjects

27 people (12 male) from 22-50 years of age (average 40 years) participated in the study. All subjects had normal or corrected to normal vision with no color blindness, were

---

[1] We placed the first target at least 31 lines from the start, ensuring that it was not visible on the first page; the time to acquire this initial target was not part of the test.

[2] This feature might add lag and thus interfere with optimal performance of the ScrollPoint or other devices (Shumin Zhai, personal communication).

right handed, and used the mouse in the right hand. Subjects had no prior experience using a mouse-integrated wheel or ScrollPoint[3]. To limit each session to 90 minutes, we used a between-subjects design where most subjects tried two devices for the task (five had enough time to use a third). Nine subjects used the ScrollPoint; 12 used Accel W1; nine used Accel W3; and 14 used the Standard Wheel. In addition, 14 subjects used a left-handed touchpad scrolling device [3], which we omit due to limited space.

Experimental Design

The design of the experiment crossed *Device* x scrolling Distance (*D*) x target Width (*W*, the tolerance). However, these were not completely counterbalanced because certain *D* and *W* interactions are meaningless (e.g., *W*=18 and *D*=6 is ill-defined because both targets fall within the tolerance without moving!). See Table 1 for a listing of all distances and widths tested. Since *Device* was not fully counterbalanced with subject (see above), we counterbalanced the order of devices across subjects using a Latin Square. Hence our analysis treats *Device* as a between subjects variable, and *Subject* as a random variable (we do not report interactions with *Subject* in our analysis, as these simply reflect the effects of individual differences).

	Width		
Distance	6	18	30
6	*		
12	*		
24	*	*	
48	*	*	*
96	*	*	*
192	*	*	*
384	*	*	*

Table 1. Widths (*W*) and distances (*D*) tested in our study. The shaded portion fully counterbalances *W* and *D*.

Following a block of practice trials for each device, subjects performed two blocks of test trials. Each block consisted of a trial for each of the 16 distance-width combinations in a random order. Thus, subjects performed 2 trials for each distance-width combination for each device. Each trial consisted of 10 or 20 *phases* of reciprocal movement between target lines; for distances of 24 or fewer lines there were 20 movement phases, while for distances of 48 or more lines there were only 10 phases, to limit the total time required to complete the experiment. Thus each subject performed 2 blocks of 16 *D/W* combinations, 12 with 10 phases and 4 with 20 phases, for a total of 400 individual scrolling movements per device.

Procedure

Participants read instructions describing the experimental set-up and task. The experimenter then reviewed these instructions with the subjects, introduced the input devices, and walked subjects through several practice trials. When the experimenter was satisfied that subjects understood the task and could perform it correctly, subjects completed the experiment unassisted. They were observed by the experimenter and encouraged to take frequent breaks. Participants typically spent 30 to 45 minutes using each device. Participants received a software or book gratuity.

RESULTS

We performed all data analyses on the mean values across phases for each trial. We used the means of the log-transformed response times from each phase to normalize the typical skewing associated with response time data.

Learning Effects

Before collapsing across the multiple phases of each trial, we analyzed the data for possible learning effects (Fig. 4). Repeated contrasts (in which the mean of each level is compared to that of the subsequent level) showed no significant difference after the second phase. Therefore, all analyses reported here exclude the first two phases.

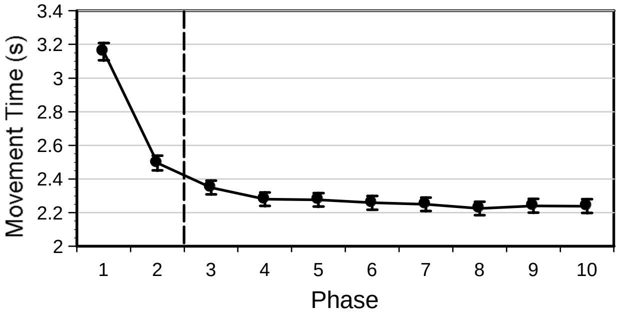


Fig. 4 Mean movement time (*MT*) for each phase. The first two phases were removed from all data analyses.

Analysis of Variance

Because *D* and *W* were not fully crossed, an analysis of variance (ANOVA) of the complete data set was not possible. Therefore, we performed a 4 (*Device*) x 3 (*W*idth) x 4 (*D*istance) ANOVA on the portion of the data that was fully crossed ($D \geq 48$ lines, the shaded portion of Table 1). This analysis revealed a significant main effect for Device, F(2,15)=15.2, p<0.001. As one would expect, movement times increased as either *W* decreased or *D* increased (i.e., as the task got more difficult: for *W*, F(2,25)=801, p<0.001; and for *D*, F(3,54)=1429, p<0.001).

Fig. 5 illustrates the aggregate means across all distance and width combinations. The two accelerated wheels appear to be faster than either the ScrollPoint or the Standard Wheel, which in turn are the same. Had we followed the example of previous studies in the literature and recorded only these overall means, we would not have found any performance differences between the Standard Wheel and the ScrollPoint.

However, because our experiment did control for distance, our ANOVA revealed a significant interaction between *Device* and *D*, F(6,45)=45.1, p<0.001. This indicates the presence of crossover effects between devices depending

---

[3] Three subjects had used an IBM Trackpoint for cursor control on a laptop; data for these participants did not significantly differ from the others.

on the scrolling distance *D*, as hypothesized back in Fig. 1. Fig. 6 (top) shows the actual crossover effect in our data.

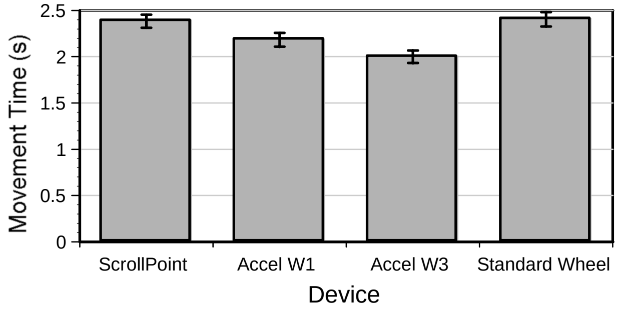


Fig. 5 Mean movement time for all devices across all widths (*W)* and distances (*D*).

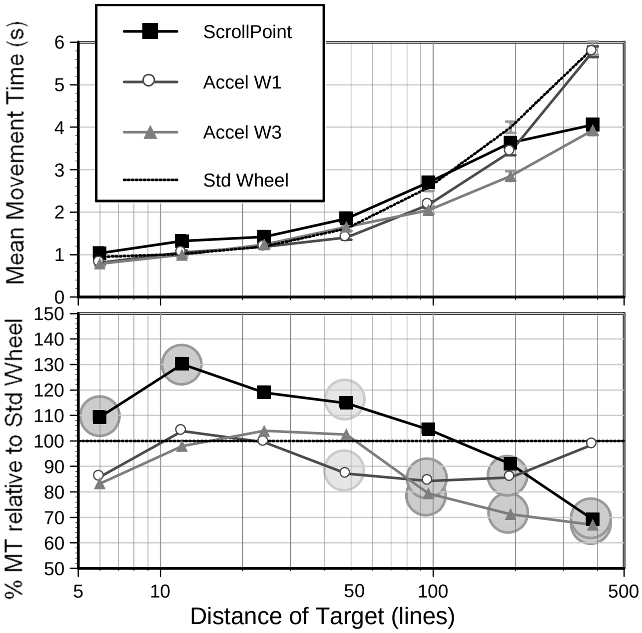


Fig. 6 *Top:* Mean movement time for all devices for each target distance (D). *Bottom: M*ovement time of each device relative to the Standard Wheel (horizontal line at 100%). Highlighted points differ significantly from Standard Wheel.

To visualize this effect, we normalized the movement time of all devices to the Standard Wheel (Fig. 6, bottom). At short distances, the wheel techniques were similar, while the ScrollPoint was significantly slower. At about 50 lines, the 1 line/notch accelerated wheel performed best, and by 100 lines, both accelerated wheels performed better than the Standard Wheel or ScrollPoint. However, by 400 lines, both the ScrollPoint and the Accel W3 were significantly faster than either the Standard Wheel or Accel W1.

We performed post hoc pair-wise comparisons between the Standard Wheel and the other devices for each level of *D*. The following data points were significantly different:

ScrollPoint: *D*=12 [t(72)=6.1, p<0.001], *D*=24 [t(118)=4.4, p<0.001], and *D*=384 [t(136)=9.2, p<0.001].

Accel W1: *D*=96 [t(154)=3.4, p<0.001], and *D*=192 [t(154)=3.6, p<0.001].

Accel W3: *D*=96 [t(136)=4.1, p<0.001], *D*=192 [t(136)=6.4, p<0.001], and *D*=384 [t(136)=9.8, p<0.001].

In addition there was a borderline effect at *D*=48 for both the ScrollPoint [t(136)=2.6, p<0.010] and Accel W1 [t(154)=2.6, p<0.011], using a Bonferonni correction of $\alpha=0.0036$ for multiple comparisons.

The ANOVA also revealed a significant interaction between *Device* and *W*, F(4,30)=2.71, p<0.049. This effect just reflects a slightly different slope for the Accel W3 mapping (Fig. 7). With other devices, it is possible that a *Device* by *W* interaction might reveal important crossover effects, so future studies should also test for this interaction.

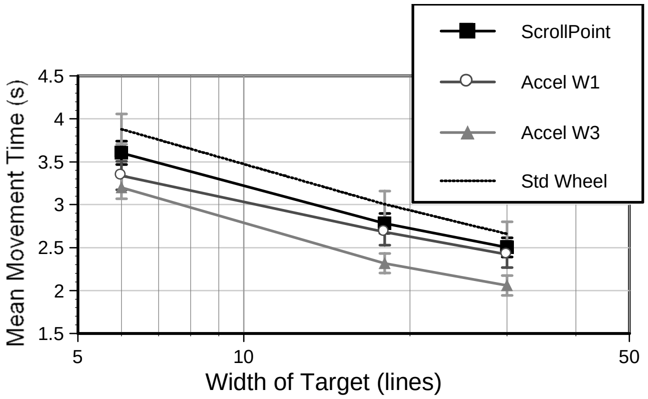


Fig. 7 Interaction between Device and W.

Does Fitts' Law Model Scrolling Movement Times?

While it was not our initial goal to model scrolling using Fitts' Law (see Equation 2), we felt it would be appropriate to see how well the model describes our data. Fitts' Law has held for a remarkably wide range of movement tasks, and we are aware of no *a priori* evidence that Fitts' Law cannot model scrolling movements. As such, the fit of Fitts' Law to this task raises an important theoretical question.

In the literature, Fitts' Law is used to predict the central tendency of the data for a single device, that is, the aggregate of all movement time data to a single data point per distance/width combination. Regression of *ID* (Fitts' Index of Difficulty) against movement time (*MT*) for each device yielded a surprisingly good fit (see Table 2 and Fig. 8 below), with correlations of .90 or higher for all scrolling devices. This is solidly within the range of correlations observed in traditional Fitts' Law pointing studies [13].

	R	**R^2**	**Slope**	**Intercept (s)**	**IP (bps)**
ScrollPoint	0.97	0.94	0.84	0.42	1.19
Accel W1	0.90	0.81	1.16	-0.51	0.86
Accel W3	0.97	0.95	0.80	0.18	1.25
Wheel Std	0.94	0.88	1.25	-0.42	0.80

Table 2 Correlation and regression coefficients for *ID* against aggregate *MT* for each device.

The Index of Performance (IP) for each device ranged from 0.80 bits/sec for the Standard Wheel to 1.25 bits/sec for

Accel W3, but this difference was not statistically significant. Compared to traditional studies of pointing devices, the IP for each of our scrolling devices seems low, but this may reflect that scrolling is a difficult task, or that scrolling input devices are in need of improvement. The y-intercepts of the models for each device were close to zero, and again were well within the range of traditional Fitts' Law studies. Hence, in the absence of contradictory evidence, the data suggest that Fitts' Law accounts for the scrolling movement times in our experimental task. Of course, better models may exist, and there is no guarantee that this result will extend to other types of scrolling tasks (e.g. [18]), but it does suggest an intriguing direction for future inquiry.

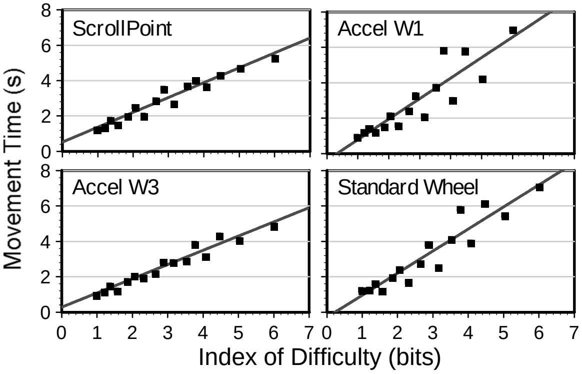


Fig. 8 Linear regression of *ID* against *MT* for each device.

Error Rates and Effective Width

Errors were possible in the study if a subject pressed the selection key while the target was not within the frame. Overall, subjects achieved a very low error rate which was close to the theoretically ideal error rate of 4% in Fitts' Law studies [13]. The error rates were ScrollPoint, 2.2%; Accel W1, 1.3%; Accel W3, 2.9%; and Standard Wheel, 3.7%.

We also recorded the endpoint coordinates of all scrolling movements (both errors and non-errors). Using the standard deviation of the endpoint coordinates, one can calculate a quantity known as the Effective Width, which is the target width that would be required to statistically correct the subject's performance to a 4% error rate [13]. Since our error rates were low and close to the theoretical optimum of 4%, we found that using the effective width in our analyses had little impact. For the Fitts' Law model, for example, the correlations for the ScrollPoint and Accel W1 were slightly better (R=0.98 and R=0.91 respectively), the correlation for Accel W3 was the same, and the correlation for the Standard Wheel was slightly worse (R=0.90).

We are also a bit distrustful of the effective width in this context. For example, the relatively coarse movement increments of the Standard Wheel mean that for some targets, the subject will nearly always "perfectly" center the line within the frame, while for others it will always be off by a few lines. It should also be noted that many researchers do not agree with the information-theoretic justification for using effective width [7]. Hence for brevity and clarity, we omit analyses based on the effective width.

QUALITATIVE RESULTS

Most subjects preferred the ScrollPoint for scrolling long distances ("it was very fast getting through long scroll distances"), but they "had trouble scrolling short lengths." However, several users commented that in practice, they would just grab the scrollbar with the mouse to move long distances. Nine of the 10 users also made at least 1 negative comment about the control or precision of the ScrollPoint: for example, users commented that it was "very hard to get a good feel for scrolling at an even speed," and that the device was "very ineffective in targeting lines." Three users liked the minimal movement required by the ScrollPoint: "my hand didn't get tired." However, two users disliked the joystick feel due to the difficulty in making fine selections.

The Standard Wheel moved predictably; as one subject put it, "I could control the speed of travel better." However, 7 of 14 subjects disliked the lack of precision, although a few preferred "scrolling through large blocks of text." Six subjects made at least 1 negative comment about fatigue or comfort with the wheel, particularly when scrolling up.

Seven subjects commented that Wheel A3 made it "very easy to scroll long distances," whereas not even one subject with the Standard Wheel made such a comment. Still, 7 subjects suggested that they preferred the ScrollPoint for long distances. Like the Standard Wheel, most subjects noticed the 3 lines/notch minimum movement, but with acceleration, 9 of the 10 subjects had a negative comment about the limited precision available. Thus, acceleration may have exacerbated the perceived control problems.

All 13 subjects who tried Wheel A1 had at least one positive comment about the increased precision. However, 10 of these subjects also had at least 1 negative comment about scrolling long distances: "It was very tedious to scroll long distances [and] required a lot of finger motion."

DISCUSSION

Here, we can further compare our study to that of Zhai et al. [18], but do recall that their experimental task differs from ours. Their documents averaged 237 lines long (at 21 pixels/line). The targets required an average of 143 lines of scrolling, whereas our study averaged 109 lines. Their screen height corresponded to W=25.6 lines. In these conditions, they found the ScrollPoint to be 29% faster than the wheel at 1 line/notch. At the most similar conditions tested in our experiment, we did not find a significant difference between the ScrollPoint and the Standard Wheel.

Jellinek & Card [11] report that increasing gain for mice does not improve performance, but rather reduces the required footprint of the mouse. Our study, which finds a performance advantage for a nonlinear control gain for the wheel, does not necessarily contradict Jellinek & Card. In fact, Jellinek & Card's argument that increased gain reduces the required footprint is consistent with our result. On a tiny device like the wheel, "reducing the footprint" is equivalent to increasing the amount of scrolling that can be achieved before the user has to release the wheel and stroke again. This reduction in total reclutching time is probably the primary source of the resulting performance advantage.

Design Implications

Our results show that the Accel W1 mapping, which has a 1 line/notch control resolution at slow speeds, performs better overall than the Standard Wheel, which uses a constant 3 line/notch mapping. Hence, in addition to improving movement times in some conditions, the acceleration function effectively enhances the control resolution of the wheel. This enables tasks such as fine placement of figures or tables using the wheel, while also enabling the device to be more effective for long range scrolling movements.

For the scrolling distances tested in this experiment, no single technique proved to be the best across all conditions. Is it possible to design a technique which is superior across the full range of distances and widths? One strategy would be to improve the ScrollPoint so that it can handle short movements more efficiently. It might also be possible to modify the wheel or optimize the acceleration algorithm to improve long distance scrolling performance, but without harming performance for short distances.

One might also attempt to combine techniques to improve performance. For example, is it possible to synthesize a combination of the Accel W1 and Accel W3 mappings? Based on our data, an algorithm that behaved like Accel W1 for distances up to about 100 lines, but like Accel W3 for longer distances, might improve overall performance. We are experimenting with an enhanced algorithm that applies this insight by keeping track of the total distance scrolled in a series of wheel movements. We have not collected formal data, but this approach seems promising.

It is known that “users tend to interact repeatedly with small clusters of information” [5], a property known as the locality of reference principle. This suggests that short scrolling movements should typically occur more frequently than long ones. However, we are not aware of any published data on the frequency distribution of scrolling movements. Our experiment points to the need for such data, as our performance data could then be weighted by a frequency distribution over distance, perhaps even on a per application or per user basis. This might be another way to optimize overall performance of scrolling techniques.

A design challenge for document navigation is its multi-scale nature [8]. Scrolling may range from a single line up to hundreds of pages in a long manuscript. By contrast, screen size limits the range of traditional pointing tasks. We believe this makes testing and controlling for a wide range of D and W even more vital when quantifying the performance of scrolling techniques.

CONCLUSIONS AND FUTURE WORK

Understanding the scrolling problem as a whole requires insight into a number of issues and problems. There are other criteria that influence performance and user acceptance, including device acquisition times, the visual diversion required to use a graphical scroll bar, or the integration of scrolling into compound tasks such as navigation plus target selection with the mouse. Our current experimental paradigm does not address these issues, but it does contribute a means to evaluate the navigational movement itself, which has presented an unsolved problem in the literature. We believe this provides a solid foundation for future studies that will further examine cognitive factors, visual attention, and other aspects of scrolling.

REFERENCES

1. Balakrishnan, R., Baudel, T., Kurtenbach, G., Fitzmaurice, G., *The Rockin'Mouse: Integral 3D Manipulation on a Plane*, CHI'97, 311-318.
2. Buxton, W., *Touch, Gesture, and Marking*, in *Readings in Human-Computer Interaction: Toward the Year 2000.* 1995, Morgan Kaufmann Publishers. p. 469-482.
3. Buxton, W., Myers, B., *A Study in Two-Handed Input*, CHI'86, 321-326.
4. Card, S., English, W., Burr, B., *Evaluation of mouse, rate-controlled isometric joystick, step keys, and text keys for text selection on a CRT*. Ergonomics, 1978. **21**: p. 601-613.
5. Card, S., Robertson, G., York, W., *The WebBook and the Web Forager*, CHI'96, 111-117.
6. Douglas, S., Kirkpatrick, A., MacKenzie, I.S., *Testing Pointing Device Performance and User Assessment with the ISO 9241, Part 9 Standard*, CHI'99, 215-222.
7. Douglas, S.A., Mithal, A.K., *Ergonomics of Computer Pointing Devices*. 1997: Springer-Verlag.
8. Guiard, Y., Baudouin-Lafon, M., Mottet, D., *Navigation as Multiscale Pointing: Extending Fitts' Model to Very High Precision Tasks*, CHI'99, 450-457.
9. Harrison, B., Fishkin, K., Gujar, A., Mochon, C., Want, R., *Squeeze Me, Hold Me, Tilt Me! An Exploration of Manipulative User Interfaces*, CHI'98.
10. Igarashi, T., Hinckley, K., *Speed-dependent Automatic Zooming for Browsing Large Documents*, UIST 2000.
11. Jellinek, H., Card, S., *Powermice and User Performance*, CHI'90, 213-220.
12. Kabbash, P., Buxton, W., *The "Prince" Technique: Fitts' Law and Selection Using Area Cursors*, CHI'95, 273-279.
13. MacKenzie, I.S., *Fitts' law as a research and design tool in human-computer interaction.* Human-Computer Interaction, 1992. **7**: p. 91-139.
14. MacKenzie, I.S., *Movement Time Prediction in Human-Computer Interfaces*, Graphics Interface '92, 140-150.
15. Microsoft Office Keyboard, *http://www.microsoft.com/ hardware/keyboard/ok_info.asp*.
16. Rutledge, J., Selker, T., *Force-to-Motion Functions for Pointing*, IFIP Interact '90, 701-706.
17. Zhai, S., *Human Performance Evaluation of Manipulation Schemes in Virtual Environments*, IEEE VRAIS'93, 155-161.
18. Zhai, S., Smith, B.A., Selker, T., *Improving Browsing Performance: A study of four input devices for scrolling and pointing tasks*, IFIP Interact'97, 286-292.