A Human Motor Behavior Model for Direct Pointing at a Distance

Regis Kopper*, Doug A. Bowman, Mara G. Silva, Ryan P. McMahan

Department of Computer Science and Center for Human-Computer Interaction
Virginia Tech
2202 Kraft Dr., Blacksburg, Va, 24060. USA
Phone: +1 540 922 3407
Fax: +1 540 231 9218

Abstract

Models of human motor behavior are well known as an aid in the design of user interfaces (UIs). Most current models apply primarily to desktop interaction, but with the development of non-desktop UIs, new types of motor behaviors need to be modeled. *Direct Pointing at a Distance* is such a motor behavior. A model of direct pointing at a distance would be particularly useful in the comparison of different interaction techniques, because the performance of such techniques is highly dependent on user strategy, making controlled studies difficult to perform. Inspired by Fitts' law, we studied four possible models and concluded that movement time for a direct pointing task is best described as a function of the angular amplitude of movement and the angular size of the target. Contrary to Fitts' law, our model shows that the angular size has a much larger effect on movement time than the angular amplitude and that the growth in the difficulty of the tasks is quadratic, rather

*Corresponding author

Email addresses: kopper@vt.edu (Regis Kopper), bowman@vt.edu (Doug A. Bowman), mara@vt.edu (Mara G. Silva), rymcmaha@vt.edu (Ryan P. McMahan)

then linear. We estimated the model's parameters experimentally with a correlation coefficient of 96%.

Key words: HCI models of human motor behavior, Fitts' law, direct pointing at a distance

---

1. Introduction

It has been argued that human-computer interaction (HCI) problems and challenges should be reduced to well-formed and rigid models in order to have a hard science of interface design (Newell and Card, 1985). In reality, very few models are robust enough that they can be used in such a way. Models of human motor behavior, however, have proven to be one exception.

Most motor human behavior models used in HCI, such as Fitts' law (Fitts, 1954) and the Law of Steering (Accot and Zhai, 1997), provide guidelines for the design of desktop interfaces, and predict time and difficulty of mouse-related tasks. When it comes to novel interfaces that require other types of input, there is little development of human motor behavior models – a model for direct interaction with volumetric displays (Grossman and Balakrishnan, 2004) and a model of raycasting in a virtual environment (Wingrave and Bowman, 2005) are notable exceptions.

Recently, many interactive systems allow the user to interact while standing up and walking around, by pointing his hand or input device directly at the display from a distance. We call this type of input *direct pointing at a distance* or, in short, *direct pointing*, and distinguish it from other methods of indicating a location on a display, such as direct touching using touch screens or indirect cursor control using a mouse.

Several existing systems and research projects use some form of direct pointing to control the cursor. Examples include home entertainment system control (Patel and Abowd, 2003) and laser pointer interaction (Olsen, Jr. and Nielsen, 2001). In fact, direct pointing interaction has made its way into consumer products with the Nintendo® Wii™, which uses direct pointing to control an on-screen cursor. Interaction with large high-resolution displays can also benefit from direct pointing (Forlines et al., 2005; Malik et al., 2005; Matveyev and Göbel, 2003; Vogel and Balakrishnan, 2005).

One important aspect of direct pointing is that it has inherent precision issues, due to hand jitter, lack of a supporting surface, and the difficulty of acquiring small targets from a remote position. To accomodate for such lack of precision, direct pointing techniques often include enhancements to allow higher precision. Enhancements may include input filtering (Vogel and Balakrishnan, 2005), area cursors rather than point cursors (Tse et al., 2007), the ability to zoom in on a particular area of the display (Forlines et al., 2005), or the ability to control cursor speed (Vogel and Balakrishnan, 2005).

Because of the complexity added by these enhancements, a user may develop several different strategies to achieve the same goal. The user must decide when and where to walk, when and where to zoom in, or when and how to change the cursor speed. This dependence on strategy makes it difficult to empirically compare direct pointing techniques with traditional techniques or with each other. Performance will depend on whether the user chose the best strategy, and individual differences may be larg.e

A predictive, analytical model for direct pointing could be used to evalu-

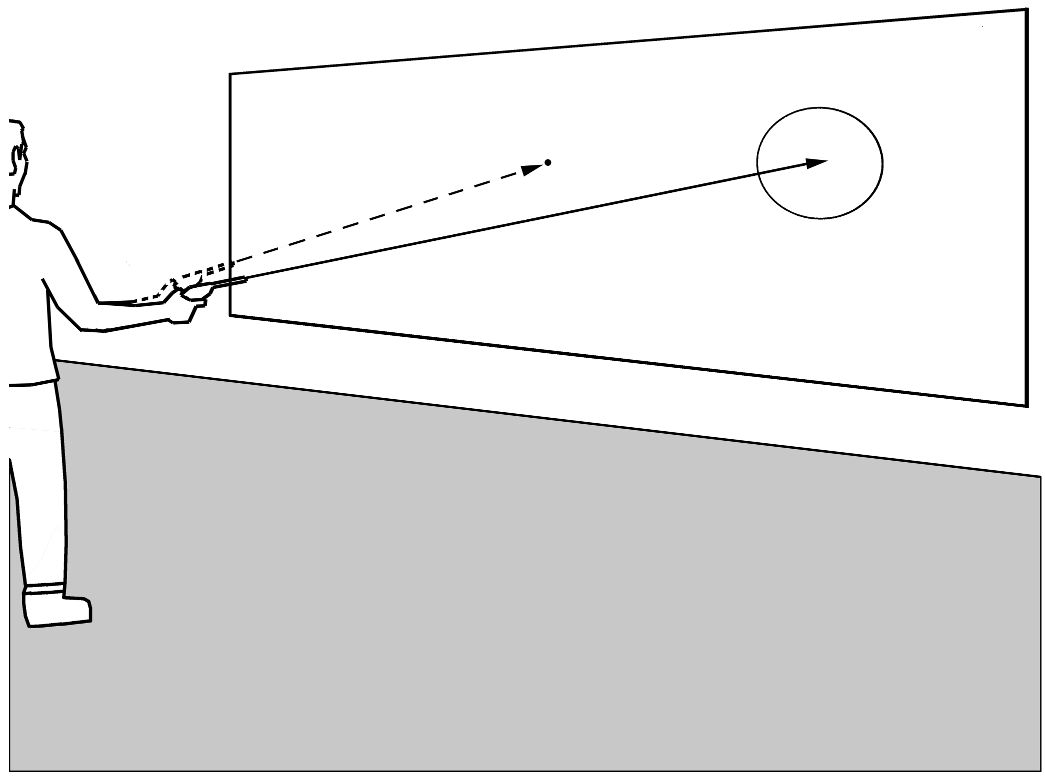

Figure 1: Direct pointing movements require arm and wrist rotations.

ate direct pointing interaction techniques, which would allow us to overcome the limitations of empirical studies. With such a model, we would be able to predict user performance in the presence of various strategies. An analytical model could also offer guidelines for the design of techniques that afford the use of good strategies, because we would know the most critical factors that affect performance on direct pointing tasks, and would design techniques accordingly.

In this paper, we present an experiment that led to a model of human behavior, inspired by Fitts' law (Fitts, 1954), for direct pointing tasks. While similar to Fitts' law, the distinctive idea that differentiates our model from most Fitts' law approaches (Accot and Zhai, 2003, 1997; Grossman and Balakrishnan, 2004; MacKenzie, 1992) is that angular target sizes and movement amplitudes, rather than linear measures, are used as input parameters for the model. Figure 1 illustrates how arm and wrist rotations are important in a direct pointing task. In addition, our model shows that angular target size is more significant than angular movement amplitude for direct pointing tasks, and that the difficulty of direct pointing tasks increases non-linearly.

In section 2 we review prior work related to our research. In section 3 we describe the different alternatives we considered for modeling direct pointing at a distance. Section 4 presents the experiment we conducted followed by its results in section 5. We then discuss some design implications derived from our findings in section 6. Finally, in section 7 we present our conclusions and consider future work.

2. Related Work

Several interaction techniques based on direct pointing at a distance have been proposed. Some of these techniques have used laser pointers as an input device (Matveyev and Göbel, 2003; Myers et al., 2002; Olsen, Jr. and Nielsen, 2001). Vogel and Balakrishnan (2005) used distant freehand pointing to interact with large displays. Jiang et al. (2006) created Direct Pointer, an interaction technique that uses a camera on a handheld device as input. We also note the similarity of direct pointing techniques to ray-casting techniques in 3D virtual environments (Bowman et al., 2004).

Most of the models that have been proposed for pointing and movement tasks are flavors or extensions of Fitts' law (Accot and Zhai, 2003, 1997; MacKenzie, 1992; Mackenzie and Buxton, 1992). The goal of these models is to predict the time to acquire a target as a function of different factors, such as the size of the target, the width of the path and the amplitude of the movement. Grossman and Balakrishnan (2005b) proposed a different approach, using a probabilistic model for the prediction of 2D target acquisition time. This research has resulted in a deep understanding of human motor behavior in pointing tasks, but it does not address our needs directly. Most prior research has focused on models in which the user either touches a target directly, or translates an input device, to cause a proportional translation of a cursor. In direct pointing, however, different types of movement are used (especially involving wrist rotation), and both the position and orientation of the input device determine the position of the cursor on the display.

Kondraske (1994) proposed a model of direct target acquisition that used angular measures in the index of difficulty, motivated by the use of joint angles to determine end-effecter position in biomechanical modeling. Although our model also uses angular measurements, we model a different task – direct pointing at a distance.

Murata and Iwase (2001) proposed a model for pointing at large information spaces. Their work, however, focused on up-close pointing in which the users touch the target with their fingers, not on direct pointing at a distance.

Grossman and Balakrishnan (2004) extended Fitts' law for trivariate targets, by modeling human performance for selecting 3D targets in a volumetric display as a factor of the width, height and depth of the target, as well as the amplitude of the movement and the angle of selection. As in our work, the input device they used was a 6-DOF tracker. However, their technique used a one-to-one mapping of the tracker position to the 3D display area. Our techniques, on the other hand, create a ray from the 6-DOF input device and intersect that ray with a flat display to determine a cursor position.

Perhaps the most similar work to our own is by Stefels et al. (2007), who evaluated different pointing devices to be used by a surgeon in an operating room. They used Fitts' law to model performance with a regular mouse, a gyroscopic relative input device and the UI Wand, which is a direct pointing input device. In their experiment, performance with the UI wand could be modeled accurately using the standard form of Fitts' law. Although they evaluated direct pointing, they did not vary the user distance to the display, which was fixed at $1.5m$, with the user remaining seated. Our work, however, focuses on conditions in which the user is standing and may interact from different distances to the display, and we eventually found that Fitts' law was not sufficient to model these more general types of direct pointing tasks.

3. Modeling Direct Pointing at a Distance

The best-known human motor behavior model for pointing tasks is Fitts' law (Fitts, 1954). Its simplicity and robustness is perhaps the reason for its heavy use in the design of graphical user interfaces (GUIs). It can be used to show that the time to acquire a target is dependent on its size and on the amplitude of movement. Generally, it can be expressed as

$$MT = a + b \cdot ID, \tag{1}$$

where MT is the movement time to complete the task, a and b are empirically determined constants, and ID is the index of difficulty of the task, which is a function of amplitude and target size.

We hypothesize that the difficulty of tasks based on direct pointing at a distance can be modeled linearly, with the slope and intercept of the regression line being determined empirically. We further hypothesize that this model contains an ID that expresses the relationship among the parameters of the task, which, for direct pointing at a distance, include target size, path length and user distance to the display. Finally, we hypothesize that, as in Fitts' law, ID a logarithmic factor, following the transmission of information theory from the Shannon's Theorem (Shannon and Weaver, 1949).

In this section, we discuss candidate formulations of ID, and discuss potential benefits and disadvantages for each of the proposed formulations.

3.1. Original Fitts' ID

Due to the uniqueness of direct pointing at a distance, as compared to other uses of Fitts' law in HCI, we believe that the traditional ID is not adequate to model direct pointing tasks. In its most widely accepted form, the original Fitts' ID (Accot and Zhai, 1999; MacKenzie, 1992; Mackenzie and Buxton, 1992) is expressed as

$$ID = \log_2\left(\frac{A}{W} + 1\right), \tag{2}$$

where A is the movement amplitude and W is the width of the target. We believe that the user distance to the display is an important factor that needs to be absorbed in the model.

Stefels et al. (2007) found a good fit of their direct pointing data to the original Fitts' law model from a fixed distance of $1.5m$ to the display.

Knowing that the distance to the display is an important element in task performance, we propose to determine whether the original Fitts' model can be used for different distances of the user to the display and affect only the slope and intercept of the Fitts' regression line (coefficients a and b from Equation 1) without loss in the accuracy of the model.

However good the fit from the original Fitts' ID is for a given distance of the user to the display, we still believe that this factor should be incorporated in the index of difficulty of the task. We wish to use our model of direct pointing to analytically evaluate direct pointing interaction techniques and strategies. Indeed, an empirical study of such techniques (Kopper et al., 2008) showed that users physically navigate relative to the display when performing realistic tasks. It is important, then, that we be able to include the user distance to the display as a parameter when doing performance predictions.

3.2. Integrating D into Fitts' ID

For direct pointing tasks, movement is not constrained to a fixed plane, such as a table, and the physical position of the user plays an important role in the difficulty of a task. To incorporate this into the Fitts' ID, we can use the raw parameters of the task, namely the amplitude of movement (A), the width of the target (W) and the distance of the user to the display surface (D), leading to an index of difficulty expressed as

$$ID_{RAW} = \log_2\left(\frac{A \cdot D}{W^2} + 1\right). \tag{3}$$

The reason for the square of the target width in this ID is that we hypothesize that the decrease in performance as W gets smaller is approximately proportional to the decrease in performance as A gets larger, or to the decrease in performance as D gets larger. When both A and D are placed in the numerator, therefore, W^2 is required in the denominator.

ID_{RAW} is more expressive than ID (Equation 2) since it accounts for the user distance to the display. It is also straightforward, since it requires only the linear parameters of the task to predict direct pointing performance. On the other hand, ID_{RAW} may not be very generalizable. Realistically, users may stand in any position in front of the display and point in any direction to perform a task. As we can see in Figure 2, it becomes unclear which value should be used for D in situations in which distance to the initial pointing location is different from the distance to the final pointing location.

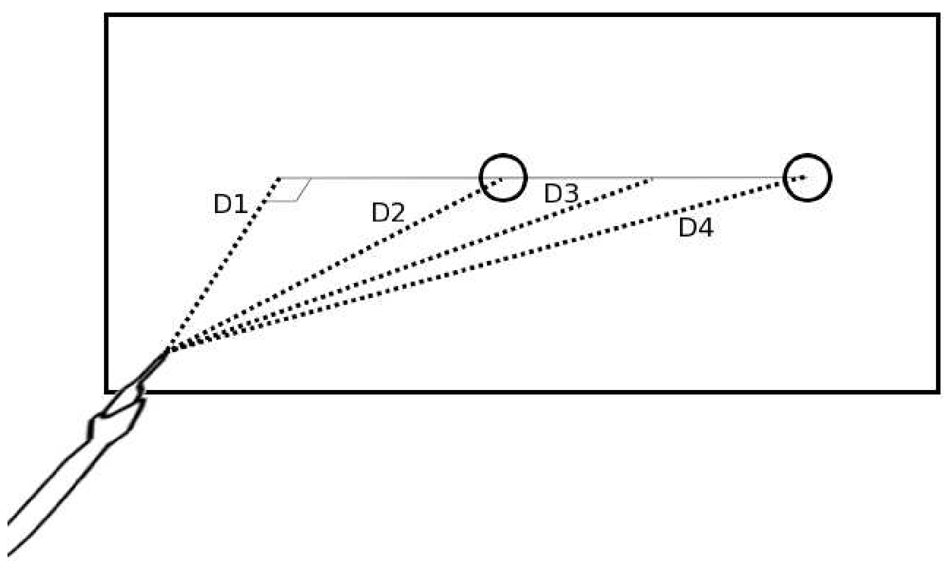

Figure 2: Ambiguity in the user distance to the display. It is not clear which distance should be used in ID_{RAW}.

We could resolve this ambiguity by using angular measurements of target size and movement amplitude in the index of difficulty for direct pointing tasks. This leads to the next proposed model.

3.3. Using angular measurements for ID

In Fitts' law terms, the amplitude of user movement in a direct pointing task decreases as the user moves away from the display because the arm or wrist rotation is smaller. For the same reason, the target width, in terms of required user movement, is smaller the farther the user is from the display. In other words, the angular amplitude of the movement (α) and the angular size of the target (ω) may be more appropriate parameters for the direct pointing model than the linear amplitude of the cursor movement or the linear width of the target object on the display. In our experimental task (see section 4), in which the user is always in the center of movement, α is formulated by

$$\alpha = 2\arctan\left(\frac{0.5A}{D}\right), \tag{4}$$

and ω is defined as

$$\omega = \arctan\left(\frac{0.5\,(A+W)}{D}\right) - \arctan\left(\frac{0.5\,(A-W)}{D}\right), \tag{5}$$

where A is the amplitude of movement, W is the width of the target, and D is the perpendicular distance from the user to the display surface. Figure 3 illustrates the analogy between the linear and the angular values.

We propose to incorporate angular measurements to the direct pointing at a distance model as a direct analogy to the original Fitts' law. This is represented as

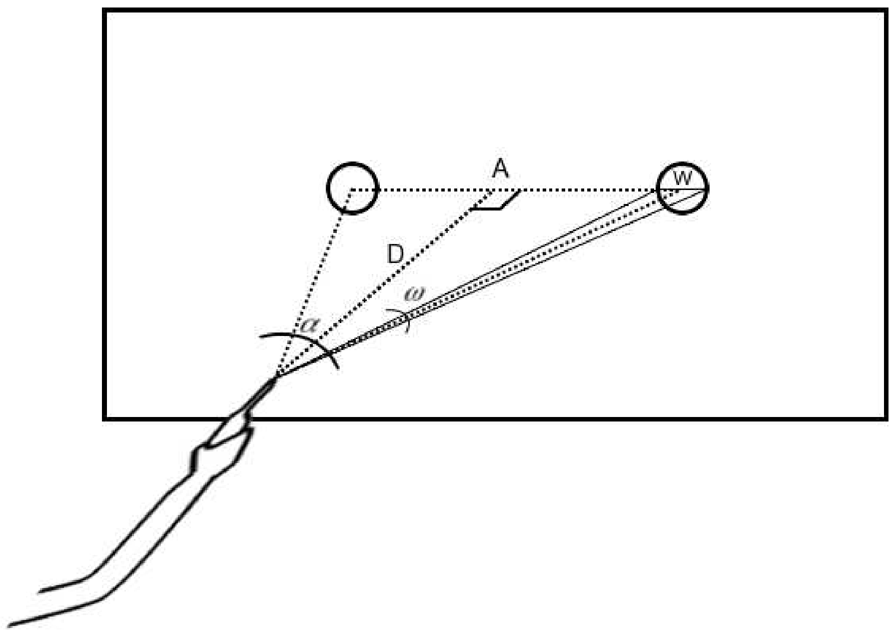

Figure 3: Relationship between α and A and between ω and W.

$$ID_{ANGULAR} = \log_2 \left(\frac{\alpha}{\omega^k} + 1 \right). \tag{6}$$

with α being the angular amplitude of movement, ω the angular width of the target and k is a constant power factor determining the relative weights of ω and α.

The reason for the constant k as a power of ω is that there is not always a linear relationship between α and ω. When using direct pointing, the user's movement consists of at least two phases: ballistic and correction (Grossman and Balakrishnan, 2005b; Liu et al., 2009; Woodworth, 1899). In the ballistic phase, the pointer moves very rapidly from one point to another using wrist rotation. These rapid movements place the pointer in the target region, and then, in the correction phase, fine-grained adjustments to acquire the target occur. In our experience with direct pointing techniques, this second phase of movement takes the majority of the time needed to complete the task (suggesting a value of k greater than 1). Further, natural hand tremor and the movement the cursor makes when a button is pressed to acquire a target – the so-called *Heisenberg effect* (Bowman et al., 2002) – add to the precision issues that make a remote target difficult and slow to acquire. We believe that different experimental settings, such as the type of input device and tracking jitter will affect the value of k.

4. Experiment

We designed and conducted an empirical study, which was similar to classical Fitts' law studies. The task we were modeling was a reciprocal selection task: the user points to and acquires, by clicking, two consecutive graphical objects.

4.1. Apparatus

We used a flat tiled display consisting of 50 NEC MultiSync LCD2080UXi monitors in a 10×5 configuration (Figure 4). Each monitor's resolution was 1600×1200 pixels, resulting in a total resolution of 16000×6000 pixels.

A wireless Iogear Phaser Mouse was used as the input device (Figure 5). A trigger-like button was used to send mouse click events, so that movement of the device due to button clicks was minimized. To enable 6-DOF input, we attached reflective markers to the wireless mouse, which were tracked by a VICON MX system with eight cameras. A dynamic recursive low pass filter (Vogel and Balakrishnan, 2005) was applied to the raw position data read from the tracker, which visibly reduced jitter without compromising the

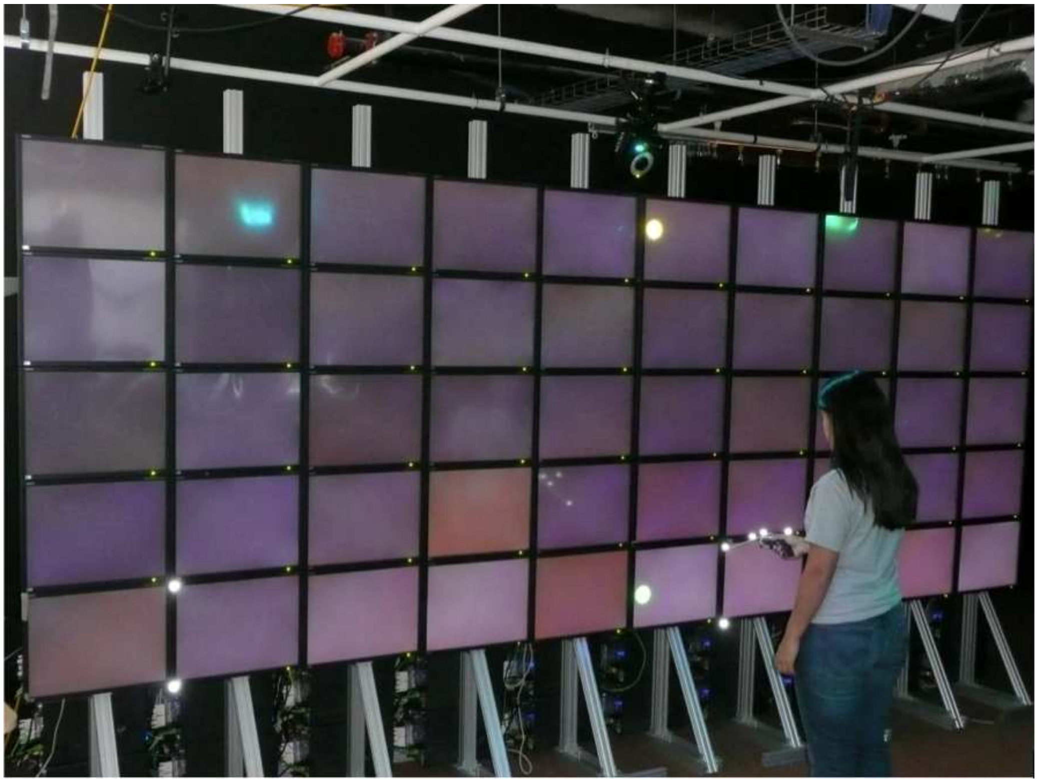

Figure 4: Large display used in our experiment.

response time. We also determined the 3D position of the display surface, enabling us to intersect a virtual ray emanating from the front of the input device with the plane of the display, and thus determining the location of the cursor. We implemented a simple application using the OpenScenegraph (www.openscenegraph.org) library. The user was presented with a black screen containing colored circular objects. Two input clients were created to handle button press events and tracker data from the input device to the application, and the data was transmitted over an Ethernet connection through sockets.

Figure 5: Input device used in the experiment.

4.2. Subjects

Twenty-one subjects (three females) were recruited from the campus community to participate in the experiment. Two were left-handed, and all subjects were instructed to hold the input device in their dominant hand. Their ages ranged from 20 to 40 years old.

4.3. Procedure

Subjects first filled in a background questionnaire and signed an informed consent form, after which they were read general instructions about the experiment.

The participants were instructed to complete tasks that consisted in selecting, via a click on the input device, two consecutive circular targets. Only two targets were shown on the screen at a time, the first was green and the second was yellow. Once the first target was selected, the targets switched colors to indicate that the subject should select the second target.

In order to get used to the interface, each subject practiced on randomly selected task combinations for five minutes. Following the practice section, the users went on to the experimental session, which is detailed below.

4.4. Design

We used a factorial within-subjects design with repeated measures. There were four independent variables: the movement amplitude A (1000, 3000, 5000 pixels), the target width W (64, 128, 256 pixels), the distance of the user to the display D (1, 2, 3 meters), and the direction of movement dir (down, up, toward dominant side, toward non-dominant side). A and W are the classical Fitts' law parameters. We included D based on our experience that direct pointing increases in difficulty farther from the display. We included dir to verify the hypothesis that different movement directions do not significantly affect task completion time. To fit the data to the model, we converted all units to meters, with $1000px$ corresponding to $0.2758m$.

The user always stood on a line extending at a right angle from the center of the display, and the center of the movement was always the center point of the screen. Thus, for horizontal tasks, the movement always went across the user's body, and for vertical tasks the movement was always directly in front of the user.

The dependent variable was task completion time (denoted in the results as MT). Each trial began when the user acquired the first target, and finished when the user clicked on the second target; the experimental software measured the trial time in microseconds. Clicks outside of the target were considered errors and were recorded, but did not invalidate the trial.

During the experiment, subjects completed three sets of trials, blocked on the distance D. We counterbalanced the order of presentation of the three distances. We made this choice so that movement fatigue and delays for adjusting to the correct position would not occur. Within each block, we randomized the order of presentation of the 36 combinations of A, W and dir. Each subject performed five consecutive trials for each of the 108 combinations, totaling 540 data samples per subject.

5. Results

Trials in which a click aimed at the second target was more than half the amplitude of movement away from the center of the target were considered mistrials and removed from the data analysis. This occurred occasionally due to the sensitivity of the trigger button in the input device, and most of the misclicks occurred inside the first target. 2.5% of the trials contained misclicks and were removed.

We considered trials that were at least two standard deviations away from the mean for each condition as outliers and removed them from the data analysis. A total of 4.9% of the trials were removed as outliers.

5.1. Movement time analysis

We performed a full factorial analysis of variance (ANOVA) for task completion time.

5.1.1. Main effects

We found significant main effects for the independent variables A ($F_{107, 10413} = 233.1, p < 0.0001$), D ($F_{107,10413} = 372.4, p < 0.0001$) and W ($F_{107, 10413} = 121.3, p < 0.0001$). No significant main effect on MT was found for the independent variable dir ($F_{107,10413} = 2.11, p = 0.0958$), which suggests that a single model is sufficient to predict performance in all four directions of movement. Figure 6 shows the least square means (t-tests) plots for the significant main effects. The standard error bars represented in the graph indicate statistical significance when they do not overlap across levels, which were all significantly different from each other.

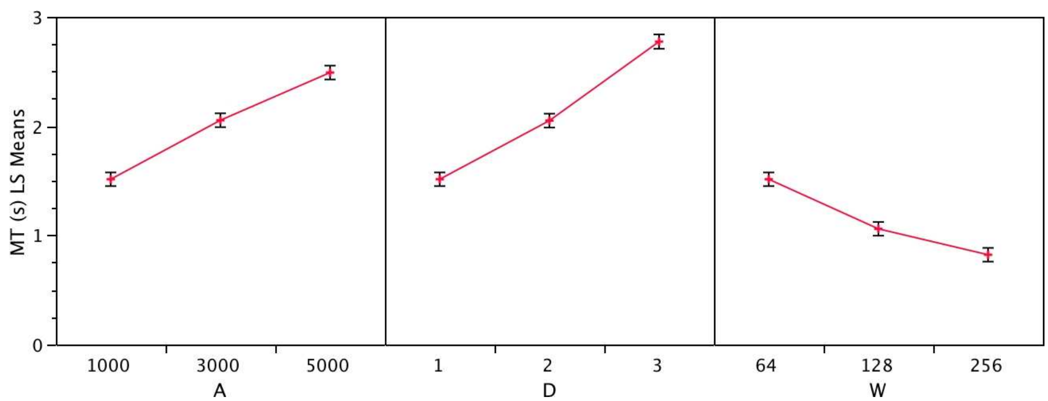

Figure 6: Least square means plots of time for significant main effects. All levels were significantly different from each other.

5.1.2. Interactions

We found a significant interaction between D and W ($F_{107,10413} = 73.09, p < 0.0001$), as seen in Figure 7. A post-hoc Tukey HSD test was performed in order to verify which levels of the interaction caused the significance. All levels of D were significantly different among each other when W was fixed at the lowest $64px$ and medium ($128px$) levels. Fixing W at $256px$, the only significant difference was between $D = 1m$ and $D = 3m$. Further, movement time was not significantly different when the user was either at the furthest distance pointing at a medium-sized target ($3m$, $128px$) or at the closes distance pointing at the smallest target ($1m$, $64px$).

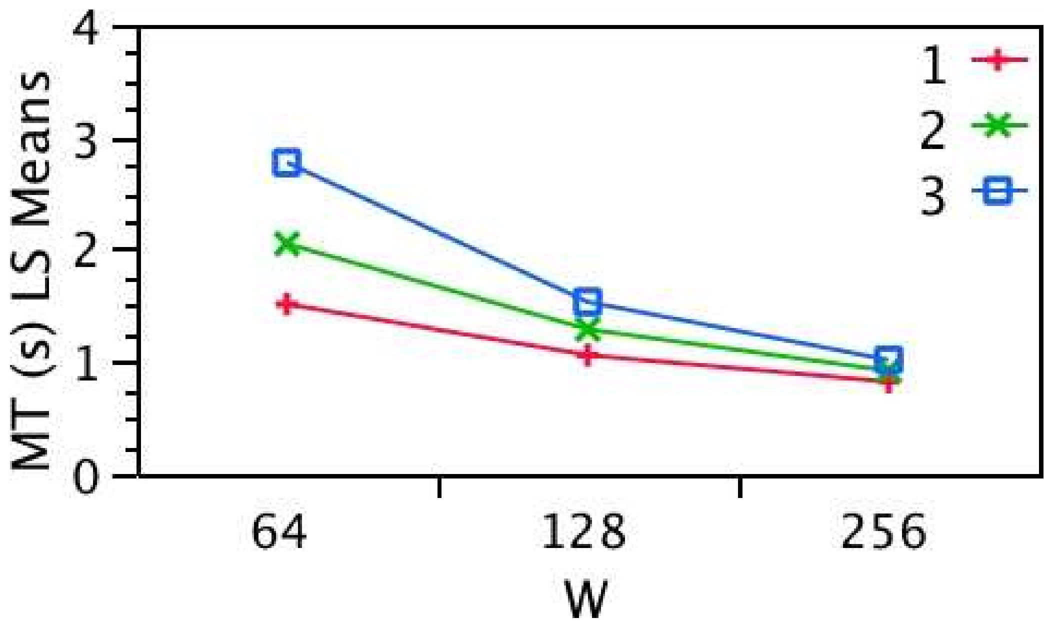

Figure 7: Interaction between D and W.

A significant, but weaker, interaction between W and A ($F_{107,10413} = 8.23, p < 0.0001$) was also detected, as shown in Figure 8. After performing a post-hoc Tukey HSD test, we found that all levels of A (1000, 3000 and $5000px$) were significantly different among each other when fixing W at each level (64, 128 and $256px$). The significantly most time consuming condition was the largest A coupled with the smallest W.

The last significant two-way interaction found was between D and A

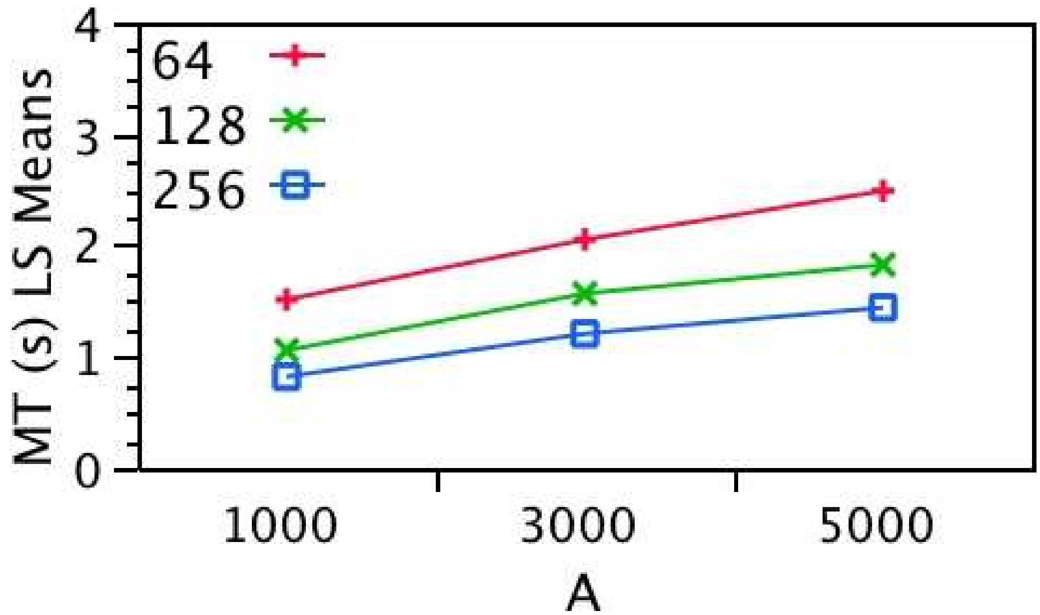

Figure 8: Interaction between W and A.

($F_{107,10413} = 2.93, p < 0.02$), as illustrated by Figure 9.

We also observed a three-way interaction between A, D and movement direction (dir) ($F_{107,10413} = 4.26, p < 0.0001$). We believe that this interaction occurred because of a glitch from the experimental setting. The Vicon tracker system used for the 6-DOF input contains near-infrared cameras that emit a bright red light. During the experimental session, some participants commented that the reflection from one of the camera's light was overlapping with the position of the target. This fact may have slowed the participants down, as they could not see the target so clearly in some conditions, and generated this interaction. Thus, we do not believe that dir plays any significant role in determining the movement time.

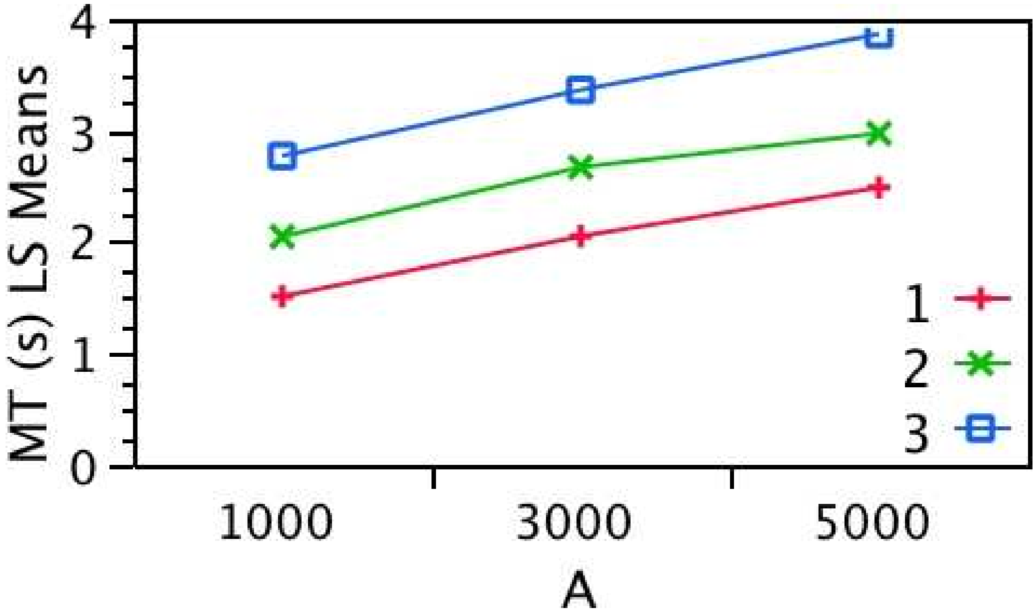

Figure 9: Interaction between D and A.

5.2. Error analysis

We counted any incorrect attempt to acquire the target as an error, but we considered the trial valid even when errors occurred. We decided not to remove trials that contained errors due to the fact that for the most difficult tasks, where D and A are high and W is low, the error rate was quite large. We can explain this due to the fact that, even with the low pass filter, hand tremor was a significant problem, especially with high values of D. The greater the length of the ray, the more sensitive will be the cursor movements. The low pass filter could have been enhanced to remove more jitter, but increasing its effect too much would lead to high movement latency. We tested several parameters for the low pass filter and found a good compromise, which reduced jitter notably without providing any perceivable latency.

We performed a full factorial analysis of variance (ANOVA) for number of errors in each task.

5.2.1. Main effects

We found significant main effects on mean error rate for A ($F_{107,10413} = 18.21, p < 0.0001$), D ($F_{107,10413} = 128.3, p < 0.0001$), and $W(F_{107,10413} = 8.999, p = 0.000124)$. Post-hoc Tukey HSD tests on the significant effects show that all the values of D were significantly different, with more errors at the further distances to the display. For the variables A and W, the only significantly different values were at the highest amplitude, and at the smallest width, respectively.

5.2.2. Interactions

There was a significant interaction between W and D with respect to the number of errors committed ($F_{107,10413} = 34.95, p < 0.0001$), and the post-hoc Tukey HSD test showed that all D levels were significantly different when fixing W at its lowest level ($64px$) with the further distances having more errors. When W was fixed at $128px$, the closest distance ($1m$) contained significantly less errors than the other two. There were no significant differences when W was fixed at its highest level.

Another significant interaction was observed when comparing the number of errors per trial between A and W ($F_{107,10413} = 3.64, p < 0.01$). The post-hoc Tukey HSD test showed that the greatest numbers of errors were committed in the highest amplitude ($2000px$) coupled with the smallest target width ($64px$).

Finally, a weaker but still significant interaction was found between A

and D ($F_{107,10413} = 2.77, p < 0.05$). The post-hoc test showed us that the combination that contained the significantly greatest number of errors was D set at $3m$ coupled with A set at $5000px$.

As with movement time, a three-way interaction of A, D and dir was found in respect to the number of errors ($F_{107,10413} = 3.60, p < 0.0001$). Again, we believe that this was caused by a camera light reflection overlapping the target in some conditions and may have caused users to make more mistakes.

5.3. Regression analysis

In order to find the best model of human motor behavior for direct pointing at a distance, we performed regression analysis of our experimental data using various possibilities for ID, as described in section 3.

5.3.1. Analysis based on the original Fitts' ID

As we hypothesized, the most common form of the Fitts' ID does not accurately model direct pointing at a distance. The regressed model of Equation 2 provided fit of $R^2 = 0.686$, which means that over 40% of the data points can't be explained by the model. This is an obvious result, since, as we can see in Figure 10, there are 3 distinct points for each ID value, each of which corresponds to one of the values of D.

We still need to verify if direct pointing could be modeled by the original Fitts' law if we have different a and b constants for each distance to the display. We regressed our data using the same ID from Equation 2, but this time, once for each of level of D.

In Figure 10 we can see that the variance increases as the distance to the display increases. Table 1 provides the coefficient of determination (R^2) for ID per distance to the display. For the up-close distance, the fit is very good, but it decreases rapidly as the distance gets larger. This analysis shows that the original Fitts' law is reliable only with direct pointing tasks with the user near the display, and is congruent with the findings of Stefels et al. (2007).

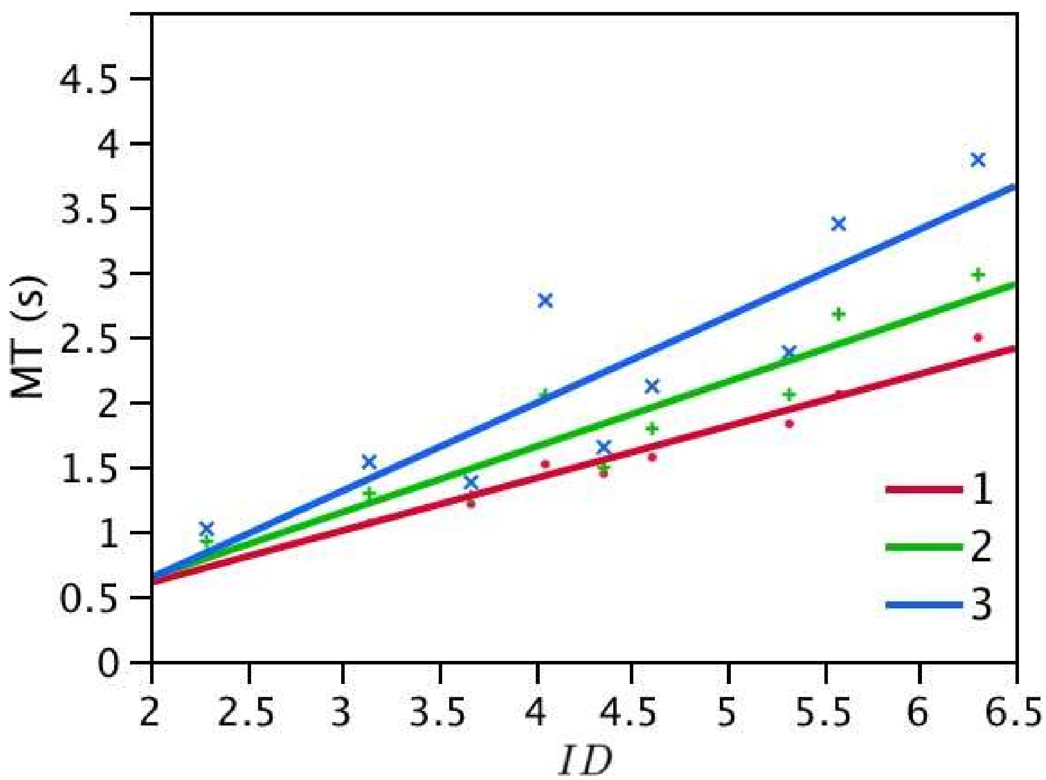

Figure 10: Fitts' law regression lines for each distance to the display based on the experimental data.

Besides lacking accuracy for higher values of D, omitting it from the index of difficulty of the task is not ideal. A more expressive model would account for for different user positions relative to the display. To achieve such expressiveness, we analyzed the data based on ID_{RAW}.

Table 1: Fit of Fitts' law for each distance to the display. a and b are the coefficients from Fitts' generic model, R^2 is the coefficient of determination and RMS is the Root Mean Square Error

$\boldsymbol{D}$(m)	$\boldsymbol{a}$	$\boldsymbol{b}$	**RMS**	R^2
1	-0.204	0.402	0.106	0.963
2	-0.362	0.502	0.267	0.864
3	-0.707	0.672	0.484	0.776

5.3.2. Analysis based on ID_{RAW}

With ID_{RAW} (Equation 3), we are able to incorporate the user distance to the display in the difficulty of the task, Figure 11 shows the linear regression of the experimental data using ID_{RAW}, with a fit of $R^2 = 0.928$.

Although the model based on ID_{RAW} fits our experimental data quite well, we believe that we can provide a more generic model if we use angular measurements in the index of difficulty. In our experiment, users always stood in the center of the movement, and we would not be able to guarantee that the same model would apply if the user stood in different positions relative to the targets on the display. By using a model that considers angular measurements, we overcome this limitation, since the angular amplitude and target width will change according to the relative position of the user to the targets in the display.

5.3.3. Analysis based on angular measurements

The model based on $ID_{ANGULAR}$ (Equation 6) results in a fit of $R^2 = 0.929$ for our experimental data, using $k = 3$, and Figure 12 shows the regres-

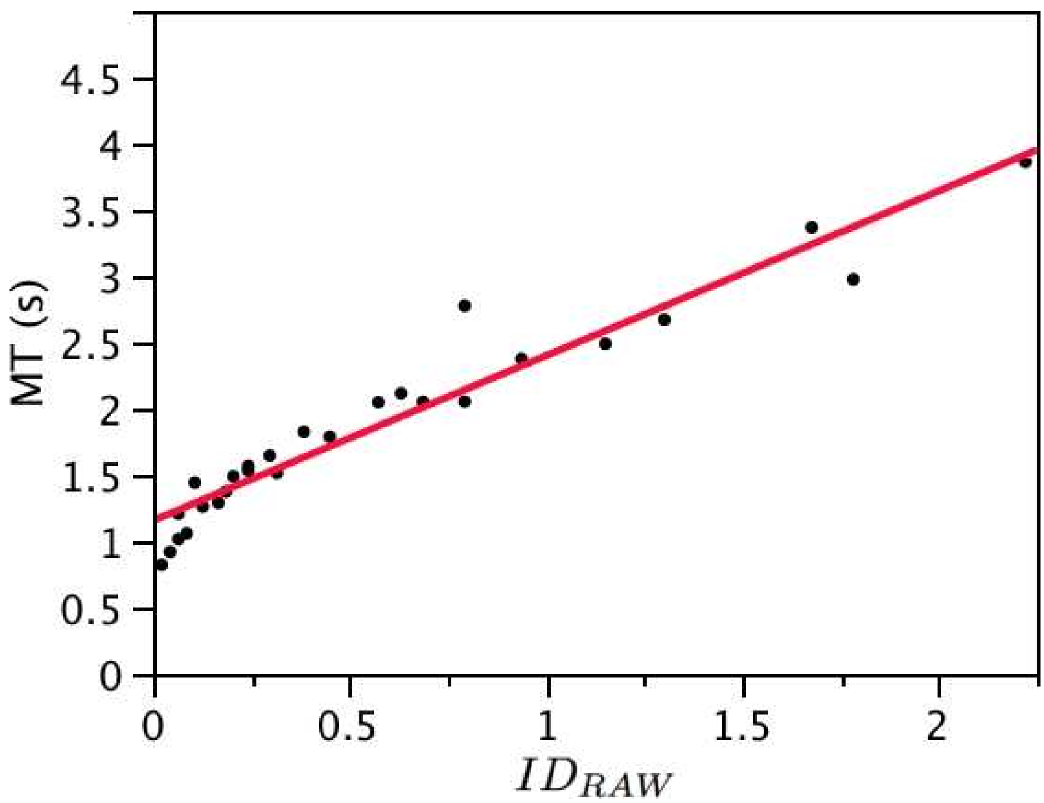

Figure 11: Regression line for ID_{RAW}.

sion line. The correlation is almost identical to the one found with ID_{RAW}, but, as argued in section 3.3, $ID_{ANGULAR}$ is more generic and expressive.

This model has a good fit and overcomes the limitations of the previous models, but it still has problems. We observed the presence of outliers at the two highest values of $ID_{ANGULAR}$, which could be suggesting an exponential trend. Such a trend makes sense, since targets with very small angular widths are very difficult to acquire. We believe that when the angular width gets extremely small, a linear increase in the time it takes to acquire the target is no longer adequate. The movement time increases exponentially, as hand

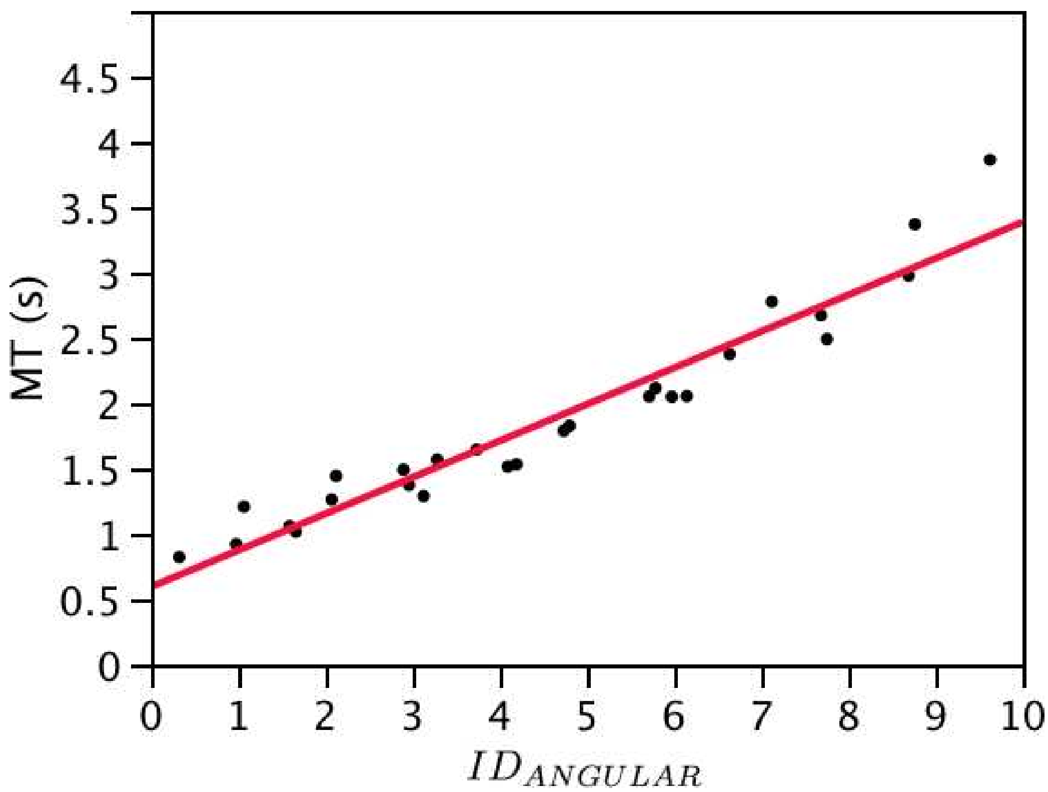

Figure 12: Regression line for $ID_{ANGULAR}$.

tremor and the Heisenberg Effect (Bowman et al., 2002) make it very difficult for the user to precisely position the cursor over the target and successfully acquire it.

5.3.4. Proposed model

In order to address the limitations of $ID_{ANGULAR}$ we examined, and ultimately adopted as our preferred model, the following *ID* for direct pointing at a distance:

$$ID_{DP} = \left[\log_2\left(\frac{\alpha}{\omega^k} + 1\right)\right]^2. \tag{7}$$

where α is the angular amplitude of the movement, and ω is the angular width of the target. To be sure that we have only one factor for the size of the target, we assume that the largest dimension of the target is parallel to the direction of movement. Using the angular width and amplitude in the ID, it is possible to account for the user distance and position relative to the display, since both α and ω will vary accordingly. Using ID_{DP} from Equation 10 we were able to fit the data to the model with a coefficient of determination (R^2) as high as 0.961 (when the value of k is 3).

It is notable that ID_{DP} consists of the square of the logarithmic factor. One could argue from intuition that the degradation of accuracy for direct pointing as the angular amplitude and size increase is more than linear. Imagine the use of a laser pointer to light up a fixed spot on a wall. When one is close to the wall, the laser dot hardly moves at all. When one steps back from the wall the laser dot begins to jitter more and more, with the position of the dot limited by a rough circle whose radius increases in inverse proportion to the distance to the wall. The amount of jitter can be expressed by the area of the circle, which increases quadratically as the radius increases linearly. This argument is supported by the results of our experiment, which show that ID_{DP} fits the data better than any of the other candidate models.

We performed regression analysis of models containing different ID_{DP} by varying the value of the constant k in Equation 10. From the ANOVA results, we saw that there is more variance in the factors related to ω than to α, so we did not expect to see good fits for k values smaller than 1. The k value that best fits the data is 3.14, but we decided to use the rounded value of 3 for the sake of simplicity to discuss our regression analysis.

The predictive model of performance for direct pointing tasks under our experimental conditions is described as

$$MT = 1.091 + 0.028 ID_{DP}, \tag{8}$$

where MT_{DP} is the movement time to complete a task and ID_{DP}, expressed in bits is

$$ID_{DP} = \left[\log_2\left(\frac{\alpha}{\omega^3} + 1\right)\right]^2. \tag{9}$$

A scatter plot with the regression line of this model is shown in Figure 13.

The fit of 96.1% and the scatter plot shown in Figure 13 were computed based on the mean for all the subjects per ID_{DP} value. This method has been used before (Accot and Zhai, 1997) and obviously provides the best fit of the data, since variance among subjects is not considered. To show that, even with the high variance from individual differences, we still can get a reasonable fit for the model when one data point per subject is used, we provide Figure 14. The fit of the model, for one point per subject per ID_{DP} value is $R^2 = 0.864$. In the figure, one can see how the variance increases with ID_{DP}, showing that the most difficult tasks depend more on individual differences, such as concentration and hand steadiness, than the easier ones.

These findings offer evidence that the model using ID_{DP} (Equation 10) accurately models direct pointing tasks. In addition, ID_{DP} is supported by the interactions among A, D and W (sections 5.1.2 and 5.2.2).

With respect to movement time, we see that the interaction between D and W suggests that, as predicted by ID_{DP} (Equation 10), the angular

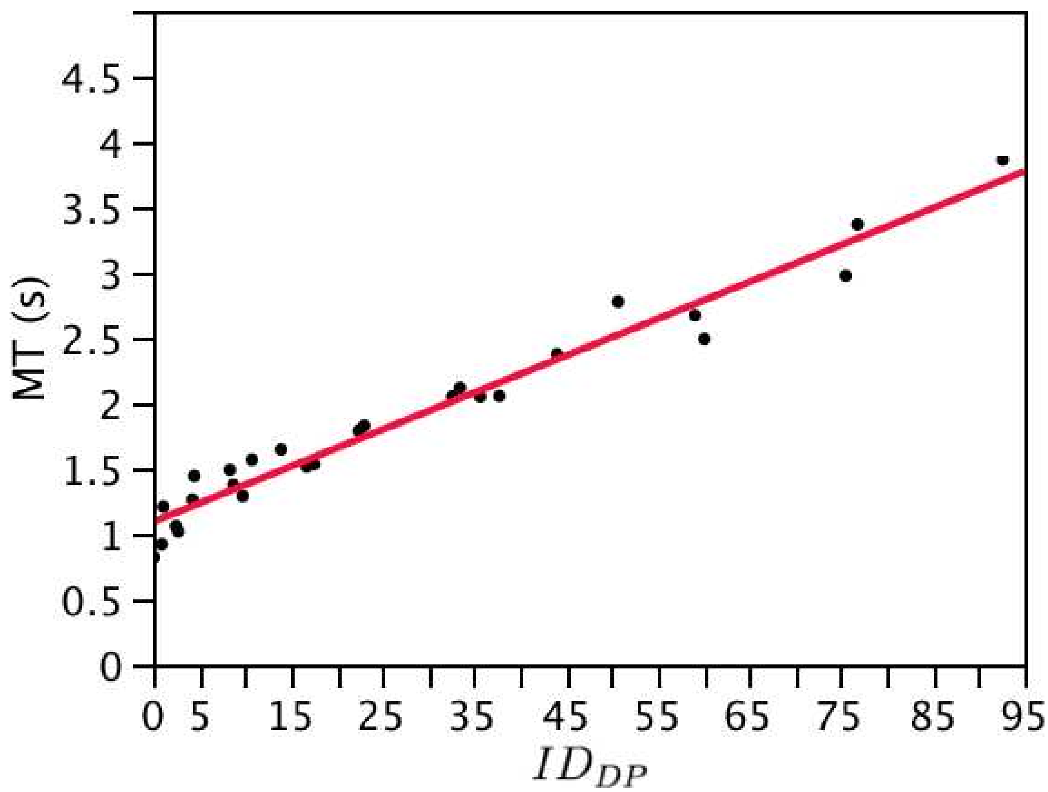

Figure 13: The scatter plot with the regression line for the fit of the model shown in Equation 8, with $R^2 = 0.961$.

width of the target (ω, Equation 5), which is a function of D and W, is a determining factor in the time it takes to complete the task.

The interaction of W with A is captured by ID_{DP} in that the user was always positioned in the center of the movement, so that the larger amplitudes cause ω to be narrower, increasing the level of difficulty of the task. Similarly, the significantly fastest condition was when the largest W ($256px$) was coupled with the smallest A (1000), which denotes the widest ω, and lowest ID_{DP}, in our proposed model. Interestingly, there was no significant difference in the movement time with the combinations of lowest W and A,

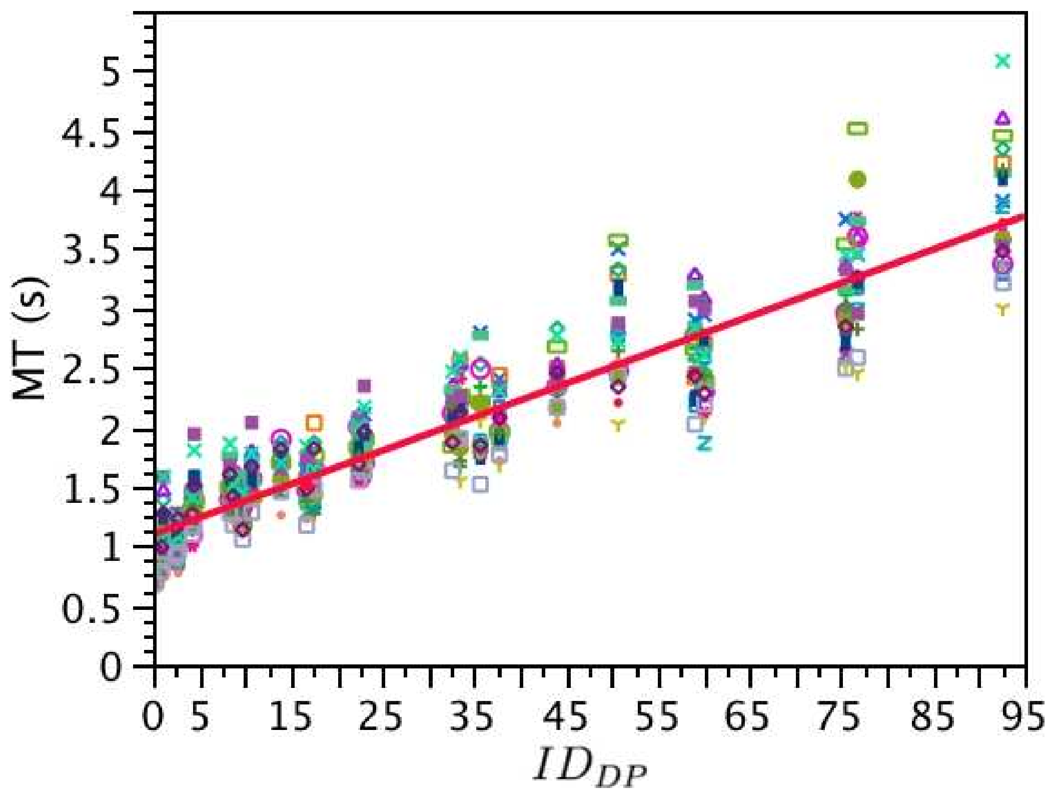

Figure 14: Scatter plot of the fit of the model with one point per subject per ID_{DP} value, with $R^2 = 0.864$. Each subject is represented by a different marker/color combination.

medium W and A, and highest W and A. Again, this indicates that the angular dimensions are appropriate, since the effective angular width of the target is smaller with the highest amplitudes, so that one factor compensates for the other in the difficulty of the task.

The interaction between D and A is reflected in ID_{DP} in the angular amplitude of movement (α, Equation 4), which is a function of the interacting factors. A post-hoc Tukey HSD test shows that when fixing D at each level, the time is significantly longer the higher the amplitude, which is congruent with the prediction of α in ID_{DP}.

Comparing the interaction between D and W with the one between D and A, we see that the former has a much stronger F statistic than the latter. This indicates that the variance for the interaction that is expressed by ω in ID_{DP} (Equation 10) is much greater than the one for the interaction expressed by α, which supports a value of k greater than 1 in Equation 10.

With regards to the number of errors, the interaction between W and D is consistent with the hypothesis that the smaller the ω, the more difficult the task is, as the biggest differences are in the lower level of W ($64px$).

The interaction of A with W supports ID_{DP}, since ω is narrowest with the high amplitudes and small target widths, prompting users to make more errors.

The weaker interaction between A and D, showing that more errors occurred with the largest D combined with the largest A is not immediately obvious in the support of ID_{DP}, since α is highest with a large amplitude and small distance. However, the task difficulty is much more strongly related to ω, which is smallest with large distances and, indirectly, amplitudes.

6. Design Implications

The model of direct pointing at a distance presented in section 5 is not only of theoretical, but also practical significance. The index of difficulty ID_{DP} has several important implications for the design of UIs that involve direct pointing.

First, and most obviously, our model indicates that the angular measurements of the target size and movement amplitude are the critical factors in task performance, rather than the linear measurements. In other words, the distance of the user from the target is highly significant. A target that may seem large when the user is directly in front of it may actually be quite difficult to acquire from a large distance (e.g., if the user walks backward, or if the target and user are at opposite ends of the display).

Designers could account for this in a number of ways. The entire UI could be designed to be usable from an assumed maximum distance, which would result in uniformly large targets; however, this approach is not likely to be practical for many applications. Alternatively, the size of targets, granularity of interaction, and/or layout of the UI could adapt based on the user's distance from the display (Peck et al., 2009). The distance from the display alone, however, does not determine the difficulty of direct pointing. A target at the opposite end of the display may subtend a very small angle when the user is near to the display, and of course the angular amplitude of movements will increase as the user moves closer to the display. An adaptive UI could therefore take as input both the user's position relative to the display and the area of the display to which the user is pointing. UIs that allow for interaction in the region of the display nearest to the user (e.g., pop-up menus instead of fixed-location pull-down menus) will also tend to reduce the angular amplitude and target size.

Second, our model clearly shows that angular target size has more influence on the difficulty of direct pointing tasks than angular amplitude. While the value of k may not be as high as three for all direct pointing setups, we expect that it will always be higher than one, because of the ease and speed with which users rotate their wrists in the ballistic phase of the movement, as compared to the great difficulty of holding the input device steady and making precise movements in the correction phase. This disproportionate influence of amplitude and size is the most important distinction, in our experience, between pointing movements based on translation across a supporting surface and pointing movements based on rotation in free space. Since there is a tradeoff between angular size and angular amplitude as the user moves closer to or farther from the display, it is very useful to understand that target size should be weighted more heavily.

Thus, increasing the size of targets should be the primary concern of the UI designer. Because of limited screen space, layout concerns, or aesthetic considerations, this will not be possible to achieve directly in most cases. Fortunately, there is a wide variety of techniques that can increase the effective target size without simply increasing the scale of the entire UI. Such techniques include utilizing the edge of the display for targets (cf. the Macintosh menu bar), area cursors (Tse et al., 2007), bubble cursors (Grossman and Balakrishnan, 2005a), Object Pointing Guiard et al. (2004), expanding targets (McGuffin and Balakrishnan, 2002), sticky targets (Worden et al., 1997) user-controlled or automatic zooming (Forlines et al., 2005; Worden et al., 1997; Ramos et al., 2007), or automatic/manual control-display ratio adaptation (Forlines et al., 2006; Vogel and Balakrishnan, 2005). Although some of these techniques may be challenging to adapt to direct pointing tasks using 3D input devices, most of them should be applicable at least in concept. In fact, some of these approaches have already been used to design new interaction techniques for direct pointing at a distance (e.g., Vogel and Balakrishnan, 2005).

Finally, the quadratic growth of the index of difficulty in our model indicates that direct pointing tasks can get more difficult very quickly as the ratio α/ω^k grows. There comes a point, especially with very small targets from large distances, where such tasks become nearly impossible. Thus, besides designing UI layouts and interaction techniques that minimize angular amplitude and maximize angular target size, designers should also provide alternatives to direct pointing to allow users to continue interacting with some reasonable level of usability regardless of the user's position or the interface elements on the screen. These alternatives could range from discrete input such as button presses to cycle through the available targets, to keyboard shortcuts, to voice interfaces.

A prototype interaction technique developed in our laboratory serves as an illustration of a direct pointing technique for which our model would predict good performance. This technique, which we call Absolute and Relative Mapping (ARM) Ray-Casting, uses manual control of the C/D ratio to allow users to increase the effective angular width of targets as needed.

By default, ARM simply uses an absolute ray-casting technique in which the handheld input device defines a pointing ray, and the cursor appears at the intersection of this ray with the screen. When finer control is needed, the user presses a button that temporarily invokes a "precision mode" with a 10 : 1 C/D ratio, increasing the effective angular width of nearby targets by a factor of ten.

The increased C/D ratio decreases the value of ID_{DP} significantly, but users may still have trouble acquiring very small targets if they cannot perceive whether the cursor is over the target. Therefore, ARM also includes a zoom lens that appears around the cursor when the precision mode is active.

The user can control the level of zoom with a scroll wheel on the handheld pointing device.

We have performed informal tests with a prototype ARM technique, and have found that it clearly increases the ease and precision of selection and placement tasks with small targets on a large high-resolution display. In the near future we will improve its design and investigate ways to automatically infer the C/D ratio and zoom level depending on the users distance and angle to the cursor position, which represents the area of interest on the display.

7. Conclusions and Future Work

We have proposed and derived a model of human performance for direct pointing at a distance based on the results of an empirical study. The angular amplitude of movement and angular target width are the main parameters of the model, which, in our experimental setting, is expressed as

$$MT_{DP} = 1.091 + 0.028 \left[\log_2\left(\frac{\alpha}{\omega^3} + 1\right)\right]^2, \tag{10}$$

where MT_{DP} is the movement time to complete a direct pointing task, α is the angular amplitude of the movement and ω is the target's angular width.

This model can be used to analytically evaluate individual direct pointing techniques, to compare the performance of multiple techniques, and to compare direct pointing techniques to other techniques, such as mouse-based pointing (which can be modeled with the traditional Fitts' law). The model can also be used to guide the design of direct pointing interaction techniques.

We plan to follow up the experiment reported here with a study using α and ω as independent variables, along with the user's relative position to the targets. This experiment will check our assumption that the model can successfully predict time no matter the user's relative position to the targets on the display, as long as the angular target width and movement amplitude are known.

Finally, the model informs us that direct pointing interaction techniques should aim mainly at increasing targets effective angular widths. We suspect that the visual angular width of targets may also affect performance, and a study of this effect, especially in the presence of high C/D ratios with very small targets, should be performed.

References

Accot, J., Zhai, S., 1997. Beyond fitts' law: models for trajectory-based hci tasks. In: CHI '97: Proceedings of the SIGCHI conference on Human factors in computing systems. ACM Press, New York, NY, USA, pp. 295–302.

Accot, J., Zhai, S., 1999. Performance evaluation of input devices in trajectory-based tasks: an application of the steering law. In: CHI '99: Proceedings of the SIGCHI conference on Human factors in computing systems. ACM, New York, NY, USA, pp. 466–472.

Accot, J., Zhai, S., 2003. Refining fitts' law models for bivariate pointing. In: CHI '03: Proceedings of the SIGCHI conference on Human factors in computing systems. ACM Press, New York, NY, USA, pp. 193–200.

Bowman, D., Wingrave, C., Campbell, J., Ly, V., Rhoton, C., 2002. Novel uses of pinch gloves^{TM} for virtual environment interaction techniques. Virtual Reality 6 (3), 122–129.

Bowman, D. A., Kruijff, E., Jr., J. J. L., Poupyrev, I., 2004. 3D User Interfaces: Theory and Practice. Addison-Wesley.

Fitts, P. M., 1954. The information capacity of the human motor system in controlling the amplitude of movement. Journal of experimental psychology 47, 381–391.

Forlines, C., Balakrishnan, R., Beardsley, P., van Baar, J., Raskar, R., 2005. Zoom-and-pick: facilitating visual zooming and precision pointing with interactive handheld projectors. In: UIST '05: Proceedings of the 18th annual ACM symposium on User interface software and technology. ACM Press, New York, NY, USA, pp. 73–82.

Forlines, C., Vogel, D., Balakrishnan, R., 2006. Hybridpointing: fluid switching between absolute and relative pointing with a direct input device. In: UIST '06: Proceedings of the 19th annual ACM symposium on User interface software and technology. ACM, New York, NY, USA, pp. 211–220.

Grossman, T., Balakrishnan, R., 2004. Pointing at trivariate targets in 3d environments. In: CHI '04: Proceedings of the SIGCHI conference on Human factors in computing systems. ACM Press, New York, NY, USA, pp. 447–454.

Grossman, T., Balakrishnan, R., 2005a. The bubble cursor: enhancing target acquisition by dynamic resizing of the cursor's activation area. In: CHI '05:

Proceedings of the SIGCHI conference on Human factors in computing systems. ACM Press, New York, NY, USA, pp. 281–290.

Grossman, T., Balakrishnan, R., 2005b. A probabilistic approach to modeling two-dimensional pointing. ACM Trans. Comput.-Hum. Interact. 12 (3), 435–459.

Guiard, Y., Blanch, R., Beaudouin-Lafon, M., 2004. Object pointing: a complement to bitmap pointing in guis. In: GI '04: Proceedings of the 2004 conference on Graphics interface. Canadian Human-Computer Communications Society, School of Computer Science, University of Waterloo, Waterloo, Ontario, Canada, pp. 9–16.

Jiang, H., Ofek, E., Moraveji, N., Shi, Y., 2006. Direct pointer: direct manipulation for large-display interaction using handheld cameras. In: CHI '06: Proceedings of the SIGCHI conference on Human Factors in computing systems. ACM Press, New York, NY, USA, pp. 1107–1110.

Kondraske, G., Nov 1994. An angular motion fitt's law for human performance modeling and prediction. Engineering in Medicine and Biology Society, 1994. Engineering Advances: New Opportunities for Biomedical Engineers. Proceedings of the 16th Annual International Conference of the IEEE, 307–308 vol.1.

Kopper, R., Silva, M. G., McMahan, R. P., Bowman, D. A., 2008. Increasing the precision of distant pointing for large high-resolution displays. Tech. Rep. TR-08-17, Virginia Tech.

Liu, L., Nieuwenhuizen, C., Martens, J., 2009. Comparing aimed movements in the real world and in virtual reality. In: Proceedings of the 2009 IEEE Virtual Reality Conference. IEEE Computer Society Washington, DC, USA, pp. 219–222.

MacKenzie, I. S., 1992. Fitts' law as a research and design tool in human-computer interaction. Human-Computer Interaction 7, 91–139.

Mackenzie, S. I., Buxton, W., 1992. Extending fitts' law to two-dimensional tasks. In: CHI '92: Proceedings of the SIGCHI conference on Human factors in computing systems. ACM Press, New York, NY, USA, pp. 219–226.

Malik, S., Ranjan, A., Balakrishnan, R., 2005. Interacting with large displays from a distance with vision-tracked multi-finger gestural input. In: UIST '05: Proceedings of the 18th annual ACM symposium on User interface software and technology. ACM Press, New York, NY, USA, pp. 43–52.

Matveyev, S. V., Göbel, M., 2003. The optical tweezers: multiple-point interaction technique. In: VRST '03: Proceedings of the ACM symposium on Virtual reality software and technology. ACM Press, New York, NY, USA, pp. 184–187.

McGuffin, M., Balakrishnan, R., 2002. Acquisition of expanding targets. In: CHI '02: Proceedings of the SIGCHI conference on Human factors in computing systems. ACM, New York, NY, USA, pp. 57–64.

Murata, A., Iwase, H., 2001. Extending fitts' law to a three-dimensional pointing task. Human Movement Science 20 (6), 791–805.

Myers, B. A., Bhatnagar, R., Nichols, J., Peck, C. H., Kong, D., Miller, R., Long, A. C., 2002. Interacting at a distance: measuring the performance of laser pointers and other devices. In: CHI '02: Proceedings of the SIGCHI conference on Human factors in computing systems. ACM Press, New York, NY, USA, pp. 33–40.

Newell, A., Card, S., 1985. The prospects for psychological science in human-computer interaction. Human-Computer Interaction 1 (3), 209–242.

Olsen, Jr., D. R., Nielsen, T., 2001. Laser pointer interaction. In: CHI '01: Proceedings of the SIGCHI conference on Human factors in computing systems. ACM Press, New York, NY, USA, pp. 17–22.

Patel, S., Abowd, G., 2003. A 2-way laser-assisted selection scheme for handhelds in a physical environment. LECTURE NOTES IN COMPUTER SCIENCE, 200–207.

Peck, S. M., North, C., Bowman, D., March 2009. A multiscale interaction technique for large, high-resolution displays. In: 3D User Interfaces, 2009. 3DUI 2009. IEEE Symposium on. pp. 31–38.

Ramos, G., Cockburn, A., Balakrishnan, R., Beaudouin-Lafon, M., 2007. Pointing lenses: facilitating stylus input through visual-and motor-space magnification. In: CHI '07: Proceedings of the SIGCHI conference on Human factors in computing systems. ACM, New York, NY, USA, pp. 757–766.

Shannon, C., Weaver, W., 1949. The Mathematical Theory of Communication. University of Illinois Press, Urbana, Illinois.

Stefels, C. N., Kneissner, J., Aarnink, R. G., Kaufholz, P., Grimbergen, C. A., Dankelman, J., 2007. Equipment control in a sterile environment using the gyromouse and a new interface, the user interface (ui) wand. Minimally Invasive Therapy and Allied Technologies 16 (3), 163 – 172.
URL http://www.informaworld.com/10.1080/13645700701384157

Tse, E., Hancock, M., Greenberg, S., 2007. Speech-filtered bubble ray: Improving target acquisition on display walls. Tech. Rep. 2007-867-19, Department of Computer Science, University of Calgary.

Vogel, D., Balakrishnan, R., 2005. Distant freehand pointing and clicking on very large, high resolution displays. In: UIST '05: Proceedings of the 18th annual ACM symposium on User interface software and technology. ACM Press, New York, NY, USA, pp. 33–42.

Wingrave, C., Bowman, D., 2005. Baseline factors for raycasting selection. In: Proceedings of HCI International.

Woodworth, R., 1899. The accuracy of voluntary movement. The Macmillan Company.

Worden, A., Walker, N., Bharat, K., Hudson, S., 1997. Making computers easier for older adults to use: area cursors and sticky icons. In: CHI '97: Proceedings of the SIGCHI conference on Human factors in computing systems. ACM Press, New York, NY, USA, pp. 266–271.