2003

Effects of pictorial depth cues on size perception in a target acquisition task

Lucia Arsintescu
San Jose State University

Recommended Citation

Arsintescu, Lucia, "Effects of pictorial depth cues on size perception in a target acquisition task" (2003). *Master's Theses*. 2473.
DOI: https://doi.org/10.31979/etd.dhh3-fmf8
https://scholarworks.sjsu.edu/etd_theses/2473

EFFECTS OF PICTORIAL DEPTH CUES ON SIZE PERCEPTION IN A TARGET ACQUISITION TASK

A Thesis
Presented to
The Faculty of the Department of Psychology
San Jose State University

In Partial Fulfillment
Of the Requirements for the Degree
Master of Arts

by
Lucia Arsintescu
December 2003

APPROVED FOR THE DEPARTMENT OF PSYCHOLOGY

Dr. Kevin Jordan

Dr. Ronald Rogers

Dr. Min-Ju Liao, NASA Ames Research Center

APPROVED FOR THE UNIVERSITY

ABSTRACT

EFFECTS OF PICTORIAL DEPTH CUES ON SIZE PERCEPTION IN A TARGET ACQUISITON TASK

by Lucia Arsintescu

The ability to judge an object in terms of size with satisfactory degree of accuracy is a very important aspect of spatial performance in virtual environments and it depends to a large degree on how these environments can be created to take a more three-dimensional form. Using Fitts' law, this study examined whether perceived size and hence movement time as predicted by Fitts' law, are influenced by the depth cues (linear perspective and foreshortening) depicted on a computer display. Moreover, the movement times for the horizontal and vertical approach to the target were examined in the presence of the depth cues. Twenty-eight participants were asked to use a trackball to move a cursor into a target object as quickly and accurately as possible. The results indicated that the observers responded to the proximal properties of the stimuli. Movement time decreased as the size of the target decreased and increased as the distance to the target increased as predicted by Fitts' law although the relationship between movement time and the index of difficulty was not completely linear. The movement time for the horizontal approach was significantly faster than the movement time for the vertical approach regardless of the depth cues presence.

ACKNOWLEDGEMENTS

It is my pleasure to thank my advisor, Dr. Kevin Jordan, for all the support he gave me while working towards my M.A. degree, and for helping me to become a researcher. I especially appreciate his huge capacity of being patient and always ready to provide help and advice. I enjoyed working with him and I hope he enjoyed working with me too. I would also like to thank to my committee members, Dr. Min-Ju Liao and Dr. Ronald Rogers, for their help and support. I would also like to thank Audra Ruthruff for the help she gave me whenever I needed it. Well, nothing would have been possible without my family, especially my husband, Bogdan. I thank them and I love them very much.

TABLE OF CONTENTS

SECTION PAGE

LIST OF TABLES

LIST OF FIGURES

Introduction

The perception of depth is an important factor that determines the efficacy of many computer-based environments. Moreover, the ability to judge an object in terms of its size with satisfactory degree of accuracy is an essential aspect of spatial performance in virtual environments and it depends to a large degree on how these environments can be created to take a more three-dimensional form. Effective development and application of such systems relies, in part, on understanding the way in which depth cues are employed by users. Another factor of importance is the motor process involved; i.e., how people interact with a movement device (e.g. a mouse, trackball, isometric joystick) in order to quickly and accurately point toward an object. Fitts' law is a useful predictor of target acquisition times on movement tasks where the aim is to reach a target quickly and accurately (Jagacinski, Repperger, Moran, Ward, & Glass, 1980; MacKenzie, 1992).

Fitts' law

Fitts' law is a model of human motor performance developed by Fitts in 1954 (Fitts, 1954). The law states that the time taken to complete a movement task depends on the amplitude of the move and the width of the target where the move ends. The mathematical formula is:

$$MT = a + b\,log2(2A/W)$$

where: MT = the movement time; a and b = the regression coefficients/constants; A = the amplitude or the distance to the target, and W = the width/size of the target. The part log2 (2A/W) in the equation represents the Index of Difficulty (ID) for a motor task. Thus, MT = a + bID.

Fitts (1954) and Fitts and Peterson (1964) found that the movement time increased as the movement amplitude (A) increased and as the target width (W) decreased. Therefore, there is a positive and linear relationship between movement time and the index of difficulty (as the index of difficulty increases, the movement time increases). Although initially applied to one-dimensional movements, Fitts' law has been successfully applied to two-dimensional movement tasks (MacKenzie & Buxton, 1992; Jagacinski & Monk, 1985; Card, English, & Burr, 1978). Attempts to apply the Fitts' law to three-dimensional movements were made by Murata and Iwase (2001) without satisfactory results.

The Perceptual – Motor Process

The perceptual-motor processes can be generally defined as the specific processes by which a person interprets and integrates the perceptual information in such a way that a motor act can be performed based on that information. Researchers (MacKenzie, 1992; Graham & MacKenzie, 1996) have considered the perceptual-motor processes by probing Fitts' law in the evaluation of performance of different tasks. For example, MacKenzie (1992), using three different devices (mouse, tablet, and a trackball) examined two different tasks such as "point-select" and "drag-select". The results showed that regardless of the device used, pointing was faster than dragging. With respect to the devices used, the performance was faster when the mouse was used. Similar results were obtained by MacKenzie, Kauppinen, and Silfverberg (2001). Using different measures of performance to bring forth the differences among four different devices (mouse, trackball, joystick, and touchpad) in pointing tasks, MacKenzie, Kauppinen, and Silfverberg (2001) found that the mouse was the fastest device followed by the trackball, touchpad, and, lastly, the joystick. However, these results may be confounded in that the participants were regular users of the mouse and the devices used were products of different companies. Card, English, and Burr (1978) compared four control devices: mouse, rate-controlled isometric joystick, step keys, and text keys, and found that the mouse was the most accurate and the fastest, although it wasn't true for the smallest distances. The slowest device was the step keys. Again the results may be confounded by the fact that the participants had more experience with mouse than with the other devices. In order to attempt to generalize Fitts' law to hand movement, Jagacinski and Monk (1985) compared the joystick with the helmet-mounted sight. They found that both devices resulted in performance that could be modeled by Fitts' law and therefore Fitts' law could be applied to hand movement. Comparing six cursor control devices (the absolute touchpad, optical mouse, trackball, relative touchpad, displacement joystick, and force joystick) on a target acquisition task, Epps (1986) found that the trackball was fit best by the Fitts model.

Depth Cues

Different cues are available to assist depth perception. These cues are named depth cues and they can be classified in 2 main categories: binocular depth cues and pictorial depth cues. The pictorial depth cues are represented in the objective stimulus present on the retina and include: linear perspective, interposition, height in plane, light and shadow, relative size, textural gradient, etc. Two depth cues of particular interest are linear perspective and foreshortening. Linear perspective is a pictorial cue based on the convergence of vertical parallel lines when receding in depth. Foreshortening is a depth cue based on numerous horizontal parallel lines receding in depth. Raddatz, Uhlarik, and Jordan (2001) examined the role of linear perspective and foreshortening on size perception. Their results showed that the accuracy of perceiving the width of the object was greater than the accuracy of perceiving the height of the object and that foreshortening was found to be more significant in maintaining size constancy. Studying different depth cues, Hendrix and Barfield (1995) found that the depth judgment was improved when perspective was added.

Miller (1997) examined linear perspective as a depth cue and found that the orientation of the depth cue influences the amount of the perceived depth in that the apex-up orientation produced the greatest amount of perceived depth, while the apex-down orientation produced the least. The apex-left and apex-right orientations did not differ from each other but they did differ from the apex-up and apex-down orientations. Miller (1997) also found that the perceived size of objects in the array was greater when the orientation of the depth cue was up. Van Breda and Veltman (1998) investigated the role of perspective (depicted as a grid surface) in a target acquisition aid in a cockpit display and found that the time to perform a target acquisition task was faster when the perspective was present than when it was absent.

All these studies suggest that the performance is always better when the depth cues are present. In the present study we used linear perspective and foreshortening in order to determine whether they influenced the movement time and the perceived size in a Fitts' task. Based on the Raddatz, Uhlarik, and Jordan (2001) study, which states that foreshortening supports size constancy, we assumed that in the presence of these cues the movement time would be influenced by perceived size.

The Angle of Approach

A number of researchers have studied the effect of angle of approaching the target object in a Fitts' task. For example, when the mouse was used, Card, English, and Burr (1978) found that the movement time was not influenced by the angle of approach. However, when the joystick was used, the movement time was influenced by the angle of approach; the movement time was slower for the diagonal direction than for the vertical and horizontal directions. Similar results were found by Jagacinski and Monk (1985). MacKenzie and Buxton (1992) and Whisenand and Emurian (1996) compared three angles of approach (0°, 45°, and 90°) and eight angles of approach (0°, 45°, 90°, 135°, 180°, 225°, 270°, and 315°) respectively, by using a mouse to move the cursor toward a target object. They found that movement time was longer for diagonal axes compared to horizontal and vertical axes. Their results contradict to some extent those of Card, English, and Burr (1978) who found no effect of angle of approach for a mouse. These results may be due to the different target objects used in each experiment. Card, English, and Burr (1978) used text strings as the target object, while, MacKenzie and Buxton (1992) used a rectangular target and Whisenand and Emurian (1996) used a square target.

These studies suggest that the movement time decreases when the target is approached from horizontal or vertical axes compared to the diagonal axes. The movement time for the horizontal approach was always faster than the movement time for the vertical approach. Thus, this was an important factor to take into consideration for our study. Therefore, in the present study we asked the participants to move the cursor to the target on horizontal and vertical axes. Since we intended to determine whether the movement time was affected by the depth cues presented on a display, it was important to keep the angle of approach constant at horizontal and vertical axes. In this way we controlled for any possible confounding due to the different angles of approach. Thus, we wanted to examine whether the movement time asymmetry replicated in our study and whether it maintained across the three displays used.

The Target Shape

The shape of the target is important in a target acquisition task. The movement time in a Fitts' task may be influenced by the shape of the target. The shape of the target object varies across different studies: targets used include squares (Kantowitz & Elvers, 1988; Epps, 1986; Whisenand & Emurian, 1996, Murata, 1999), circles (Jagacinski & Monk, 1985; Graham & MacKenzie, 1995; MacKenzie, Kauppinen, & Silfverberg, 2001), and rectangles (Card, English, & Burr, 1978; Mackenzie & Buxton, 1992). There is little research which compared the shape of the target within an experiment and its effect on movement time and therefore performance. Sheikh and Hoffmann (1994) examined the effect of five target shapes (rectangle, square, diamond, circle, and equilateral triangle) on movement time. They varied the width of the target (5, 10, 20, and 40 mm) although the amplitude was kept constant (320mm) across all conditions. They found no effect of shape on movement time, but an interaction of target width and target shape. The triangle and the diamond took longer to move to for the two smallest widths than took for the rectangle, circle, and square, although all target shapes were not significantly different in movement time at any target width. This difference in movement time between the triangle and diamond, and the rectangle, circle, and square may be due to the constraints of target height for the triangle and the diamond, although this was not true for the largest widths. This suggests that the shape of the target becomes an issue when the width is small. However, it would be useful to determine whether the shape of the target is influenced by the amplitude of the movement, which was not included as a variable in this study given that the amplitude was held constant.

Whisenand and Emurian (1999) investigated two target shapes (square and circle) for three target widths (4, 8, and 16 mm), four amplitudes of movement (20, 40, 80, and 160 mm), eight angles of approach (0°, 45°, 90°, 135°, 180°, 225°, 270°, and 315°), using two types of task (drag-drop and point-select). Their results showed that the movement time was significantly faster for the square target than for the circle and for targets, circle and square, movement time decreased as the target width increased. The results also showed that the frequency of errors was higher for the circle than for the square. These results contradict those of Sheikh and Hoffmann (1994) which showed that there was no difference in movement time between square and circle. Also Whisenand and Emurian (1999) did not find an interaction of target width with target shape as Sheikh and Hoffmann (1994) did. The difference in the results may be due to the different types of task; a manual tapping task in the Sheikh and Hoffmann (1994) study and a computerized task in the Whisenand and Emurian (1999) study. Also, in the first study, the authors held the amplitude of the movement and the angle of approach constant while in the second study these variables were manipulated. Therefore, the results obtained by Whisenand

and Emurian (1999) may be confounded by the presence of the other independent variables. However, it is hard to determine the most effective shape for a target acquisition task that will increase the movement time and therefore, will enhance the performance. To eliminate any effects due to the interference between the angle of approach and the shape of the target, in the present study we used a circle as the target object.

Measurement of Target Size

The first measure of target size was defined by Fitts (1954) as the horizontal width of the target. In finding the measurement for target size for two-dimensional tasks, which should be used in human-computer interaction, MacKenzie and Buxton (1992) compared the three models of target size. The SMALLER-OF model specifies that the smaller of W and H is used for target size. The W' model specifies the width of the target size as the width of the target along an approach vector. Finally, the STATUS-QUO model specifies that W is always the horizontal extent of a target. The SMALLER-OF model requires specification of only A, W, and H in order to be applied, while the W' model requires A, W, H, THETA-H, and a geometric calculation for determining the correct substitution for W, and the STATUS-QUO model requires only A and W. The manipulated variables of the study were the angle of approach, the amplitude, the target width, and target height and the measured variables were movement time and error rate. The target object was represented by a rectangle. The authors found that the SMALLER-OF and W' models were not significantly different from each other but they were different from the STATUS-QUO model. The authors argued that W' model is attractive because it maintains the one-dimensionality of the model (as in the initial Fitts' task), but is difficult to apply because it requires many parameters (W, A, H, and the angle of approach). At the same time SMALLER-OF model is more practical because it can be applied without taking into consideration the angle of approach. However, the easiest model to apply, in general, is the STATUS-QUO model because predictions are based only on W and A. With a circle, all three models predict the same outcome.

Gillan, Holden, Adam, Rudisill, and Magee (1992) compared two types of tasks: pointing and dragging using text strings. The manipulated variables of the study were distance to the target (2, 7.5, and 13.75 cm), target width (.25, 1.0, 3.5, and 6.0 cm), angle of approach (up-to-right, up-to-left, down-to right, and down-to-left), and the type of movement (point-click and point-drag). The models of calculating the target size were H and W + H. They found that for both pointing and dragging, the height of the text influenced the movement time.

Using the theory of two-dimensional probability distribution, Murata (1999) defined the effective target width for two-dimensional pointing tasks. The manipulated variables of the study were the amplitude (500, 650, 800, and 950 pixels) and the area of the target (16129 pixel^2, 33856 pixel^2, 57600 pixel^2, and 88209 pixel^2). The angle of approach was held constant at 45° (from lower left to upper right). The dependent variable was pointing time and the target object used was a square. The effective target width was determined separately for the horizontal (x) and vertical (y) axis. Therefore the target size was determined based on the width of x and y, using the formula: $s = \min(W_{xe}, W_{ye})$, where s = size, W_{xe} = the effective target width for x axes, and W_{ye} =

the effective target width for y axes. The results of the study showed that the width of the target should be calculated based on W_{xe} and W_{ye}; although the authors argued that a higher and a more predictive performance model will need to be developed. However, it was found that the proposed model is not necessarily better than the traditional performance model.

These studies suggest that the measurement of the width becomes an issue when the targets are other than circles. For circles, the best measurement approach remains that defined by Fitts (1954) in which W and A are measured along the same axis although all models described above lead to the same predictions for a circle. This is supported by the fact that the width of the circle is the same, regardless of the angle of measurement. Therefore, using a circle in the present study eliminated the need to choose among the models of target size.

The Present Study

The purpose of the present study was to examine the perceptual-motor processes by examining the movement time and the perceived size of the target when linear perspective and foreshortening were present. Specifically, using Fitts' law we examined whether perceived size was influenced by the depth cues depicted. We tested the applicability of Fitts' law in displays containing depth cues (such as linear perspective and foreshortening). Fitts' law predicts that the movement time increases when the index of difficulty increases.

The presence of the depth cues may affect the movement time. The presence of depth influences the perceived size of the target object, as it was determined by the previous studies described earlier. In particular, we examined how the size of the object was influenced by depth by determining whether observers perceived the size proximally or distally. Proximal size is the size of an object as measured on the retina and distal size is the size of an object as measured in the real world. If the observers responded to the proximal properties of the stimuli, then we expected no difference of movement time across the three display types (linear perspective, foreshortening, and control). If the observers responded to the distal properties of the stimuli, then we expected that the movement time would be different across the three display types with the movement time faster in the presence of depth cues. This is because perceived size was expected to be influenced by linear perspective and foreshortening.

The direction of movement may influence the movement time in the presence of the depth cues. If there would have been an effect of depth on the perceived size, it would have been possible that the movement time should be less for the vertical direction since perceived size should increase. The presence of depth cues could influence the movement time if observers responded to the distal properties of the stimuli. Previous studies showed that the movement time was faster for the horizontal approach of the target compared to the vertical approach of the target. Therefore, it was of interest to examine whether the horizontal versus vertical movement time asymmetry was replicated and whether it held across all three displays (linear perspective, foreshortening, and control). It is possible that the horizontal – vertical asymmetry could not be replicated in the presence of the depth cues if there is an effect of depth. Therefore, we expected that the movement time for the vertical direction will be influenced by the perceived size and depth.

The independent variables manipulated in the study were the distance (A) from the cursor to the target object, the width (W) of the object, the direction of movement (from left to right and from down to up), and the depth cues (linear perspective, foreshortening, and control). For examining the movement time, we created 12 indices of difficulty (four target sizes and three distances). The pictorial depth cues were: linear perspective present, foreshortening present, no depth cue present/control. The dependent variable was movement time made during the target acquisition task.

Method

Participants

Twenty-eight San Jose State University students between 18-35 years of age participated in the study in order to fulfill a course requirement. The participants had normal or corrected to normal vision as determined by self report. The participants used their preferred hand in accessing the device. The Human Subjects-Institutional Review Board at San Jose State University has approved our request to use human subjects for the present study (see Appendix A).

Apparatus and Stimuli

This experiment used an IBM compatible computer with an Intel Pentium processor. The visible screen on a color CRT 17-inch monitor was 1064-786 pixels. In order to produce equivalent visual displays the experimental displays were limited to 786x786 pixels. The experiment was conducted using Word Construction Set by 3D

Nature Company and a C++ program which generated the stimuli and recorded the movement times. Cursor control and target selection were controlled by a Kensington trackball. The stimuli consisted of a cursor, a target object, and a background. The cursor had the shape of a triangle, the target object was represented by a circle, and the background was represented by linear perspective (see Figure 1), foreshortening (see Figure 2), or a control (no depth cue present).

The colors of the display were stated as follows: the linear perspective and foreshortening conditions were represented by black lines on a grey background, the control condition were represented by a white background, the cursor was white with blue margins, and the target object was red. This was based on the study of Luria, Neri, and Schlichting (1989) who found that the speed and the accuracy of the search were not influenced by different phosphors used (green, amber, yellow, red, blue, and white). The target object was represented by a circle of four different sizes and was positioned at three different distances from the cursor.

The sizes were calculated by using the classic Fitts model (W), and the distances were calculated from the starting point to the center of the target. The sizes of the target objects were 10, 18, 30, and 36 mm and the distances from the cursor to the target were 60, 170, and 250 mm. Thus, the indices of difficulty (ID = log2 (2A/W) created were: 1.74, 2.01, 2.75, 3.25, 3.51, 3.60, 3.81, 4.07, 4.25, 4.81, 5.10, and 5.66 (see Table 1).

The movement directions were from left to right and from down to up. In the linear perspective condition, several vertical lines were slanted into the horizon, and in the foreshortening condition several parallel horizontal lines appeared closer as they go to

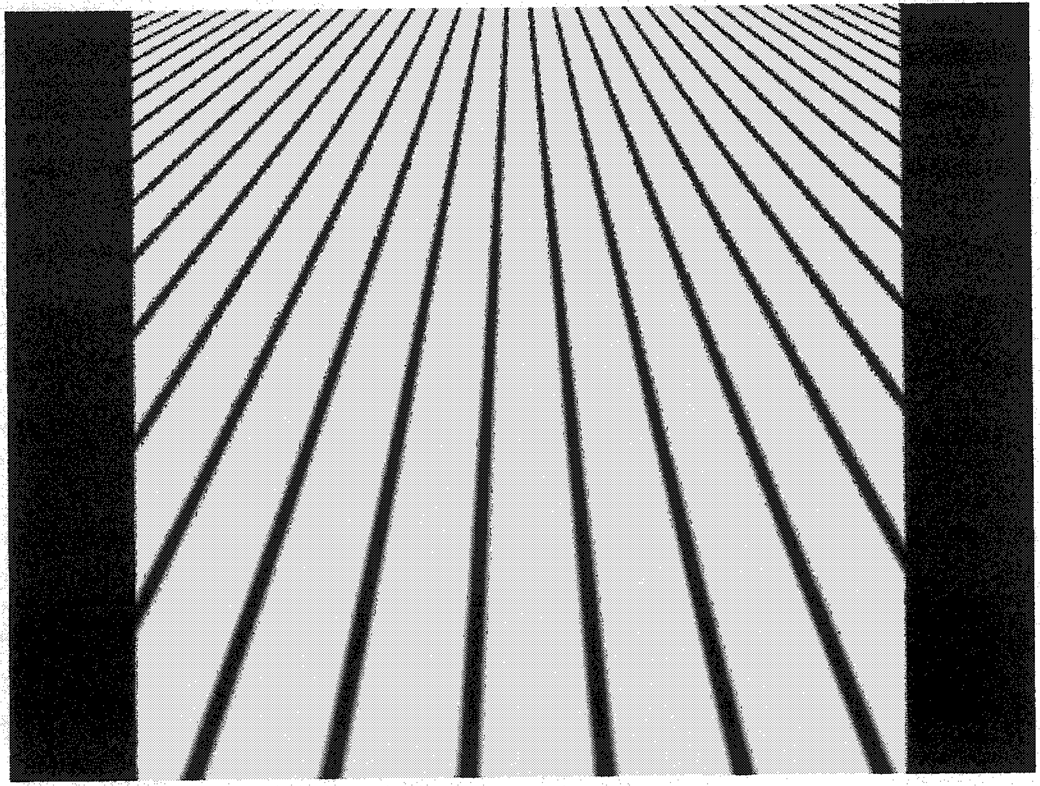


Figure 1. Linear Perspective Display.

Figure 2. Foreshortening Display

Table 1

The Indices of Difficulty (ID), Sizes (W), and Distances (A)

W	A		
	60mm	170mm	250mm
10mm	**3.60**	**5.10**	**5.66**
18mm	**2.75**	**4.25**	**4.81**
30mm	**2.01**	**3.51**	**4.07**
36mm	**1.74**	**3.25**	**3.81**

Note. The numbers in bold represent the ID for the corresponding size and distance.

the horizon. To summarize the design of the experiment, there was a 4 (target sizes: 10, 18, 30, and 36 mm) x 3 (distances: 60, 170, and 250 mm) x 3 (background: linear perspective, foreshortening, and control/no depth cue present) x 2 (directions of cursor movement: from left to right and from down to up) within subjects design. The factorial combination of these four independent variables resulted in 72 trials.

Procedure

The experiment took approximately 40 minutes. Each subject was tested individually. Before starting the experiment the participants were given a consent form to sign (see Appendix B). In the consent form was explained the risks that participants may have during the experiment and their rights with regard to the participation. After signing the consent form, participants were given instructions describing the task of the experiment (see Appendix C). The participants were asked to use the trackball to move the cursor (placed either on the left of the screen or on the bottom) into the circular target as quickly and accurately as possible. Each trial was terminated when the cursor entered the center of the target. To start the experiment and to proceed from a block of trials to another the participants were asked to press any key on the keyboard. The different display conditions were not explained to the participants. Before beginning the experiment the participants were given 12 training trials. The training trials consisted of the three background conditions, the two movement directions used in the experiment, two sizes (which were different than those used in the experiment in order to control for practice effect), and the same three distances used throughout the experiment. The practice trials were randomized across participants, thus not all participants were given the same order of trials. Participants were encouraged to ask any question related to the task during training. The training session was meant to familiarize the participants with the display and the device (trackball) and also to ensure that they understood the task requirements.

After the training session was completed, participants were informed that the experiment would begin. They were told that the experiment would involve 288 trials (four blocks of 72 trials) and that they could take a short break between the blocks if needed. The presentation order of the conditions was randomized across participants and across blocks. Thus, not all participants were given the same block first and within the blocks the participants did not received the same order of trials. The movement time for each trial was recorded on the computer disk. The movement time was calculated as the time taken to move the cursor into the target object starting at the very first second of the trial and ending when the cursor reached the target. Therefore, for each participant there were 300 data points (12 practice trials and 288 experimental trials). At the end of the experiment the participants were provided with a debriefing statement explaining the purpose of the study (see Appendix D).

Results

In order to assess the appropriateness of using Fitts' law to model the movement time (MT) data, means (*M*) were examined for movement time as a function of the index of difficulty (ID) as shown in Table 2 and Figure 3. The MT for each background (linear perspective, foreshortening, and control) was computed as a function of the index of difficulty and this is shown in Figure 4.

Table 2

Summary Statistics of Movement Time as a Function of Task Difficulty (n = 28)

	M(ms)	*SD*
$MT_{1.74}$	740.56	149.88
$MT_{2.01}$	700.72	166.47
$MT_{2.75}$	860.73	191.07
$MT_{3.25}$	1190.74	172.90
$MT_{3.51}$	1162.74	162.40
$MT_{3.60}$	1047.76	237.10
$MT_{3.81}$	1421.79	240.35
$MT_{4.07}$	1365.18	222.02
$MT_{4.25}$	1307.86	204.08
$MT_{4.81}$	1535.31	233.94
$MT_{5.10}$	1513.23	254.18
$MT_{5.66}$	1737.59	272.81

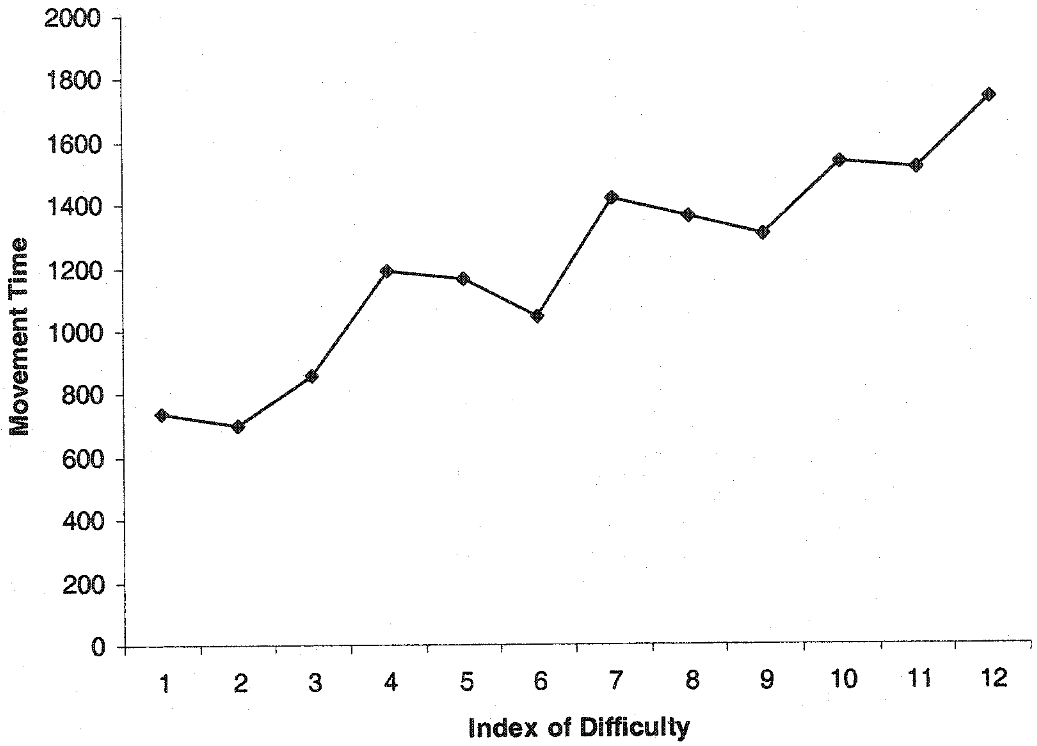


Figure 3. Movement Time as a Function of Task Difficulty. Movement time increased as the index of difficulty increased as predicted by Fitts' law.

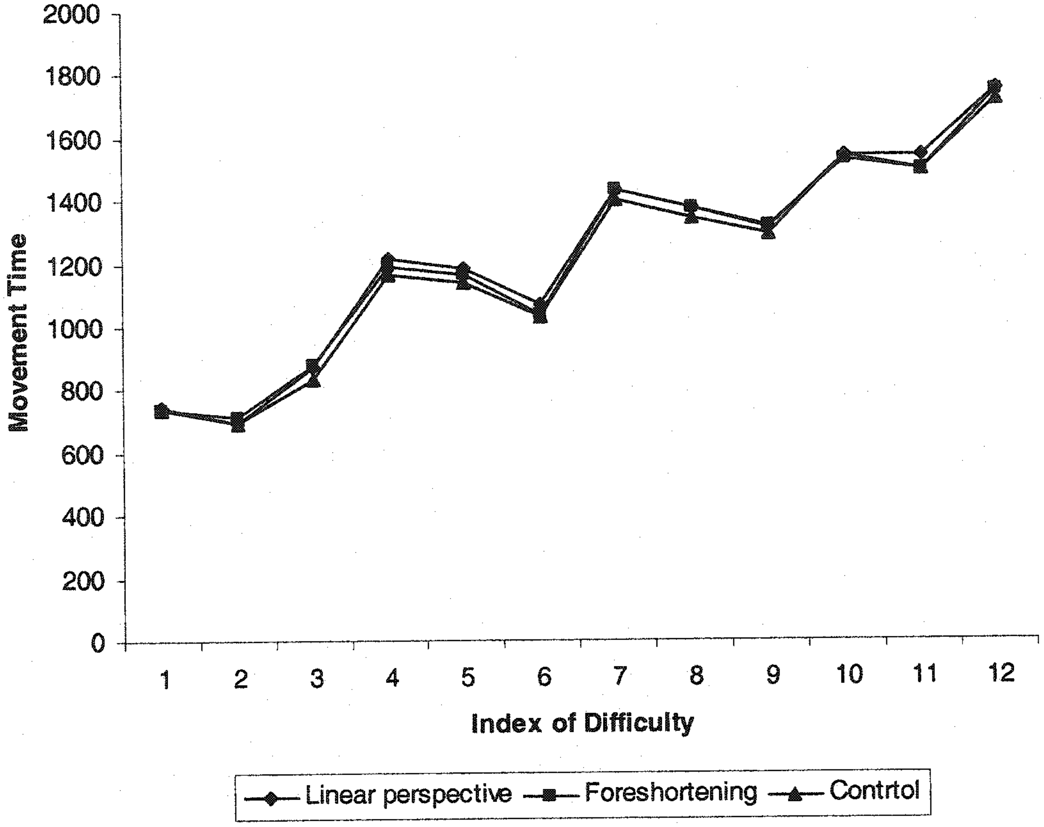
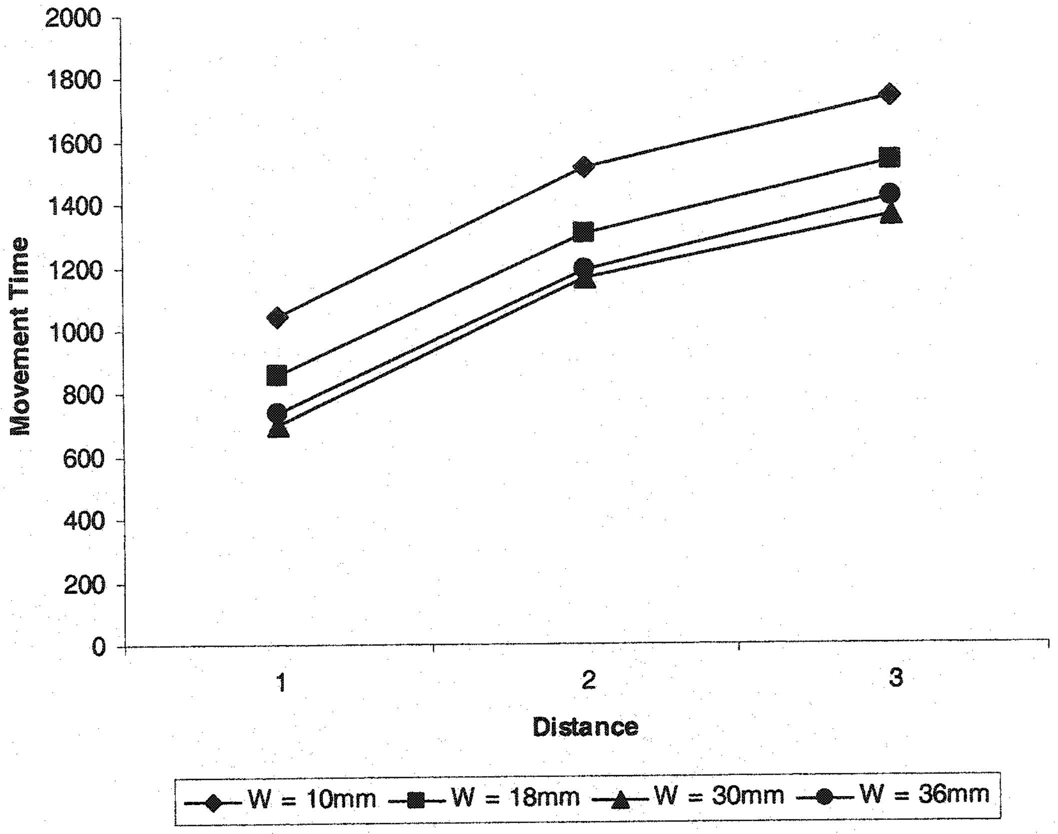


Figure 4. Movement Time for Each Background as a Function of Task Difficulty. For each background display (linear perspective, foreshortening, and control) the movement time increased as the index of difficulty increased.

Linear regression analyses using MT as the dependent variable and the ID as the independent variable was conducted to determine the linearity of the Fitts' law model across the three display types and we found $R^2 = .93$, $p < .01$ for linear perspective, $R^2 = .92$, $p < .01$ for foreshortening, and $R^2 = .93$, $p < .01$ for control. The average R^2 across the three displays was $R^2 = .93$, $p < .01$. However, as seen in Figure 4 there are clear departures from linearity in the relationship between the MT and the ID. Therefore, additional regression analyses were conducted for each target size separately. The results showed for W = 10mm $R^2 = .997$, $p < .01$; for W = 18mm $R^2 = .995$, $p < .01$; for W = 30mm $R^2 = .999$, $p < .01$; and for W = 36mm $R^2 = .994$, $p < .01$. As shown in Figure 5, the relationship between MT and distance is linear for each target size. These results suggest that the data does fit Fitts' law.

In order to determine whether the observers respond to the proximal or distal properties of the stimuli, we combined the 4 sizes and the 3 distances into the ID's (as shown in Table 1). Thus a 3 (background: linear perspective, foreshortening, no cue/control) x 2 (two directions of movement: horizontal and vertical) x 12 (ID: the 12 combinations of sizes and distances) within-subjects ANOVA was performed on movement time. There was a main effect of ID ($F(11, 297) = 368.62$, $p < .001$) and a main effect of background ($F(2, 54) = 8.09$, $p < .001$). Post hoc analyses of ID's indicated that all of them were significantly different from each other at $p < .05$, as shown in Table 2 and Figure 3. With respect to the ID's, in general, the MT increased as the ID's increased although for some ID's the MT was longer than the MT for the previous ID



Distance: 1 = 60mm; 2 = 170mm; 3 = 250mm

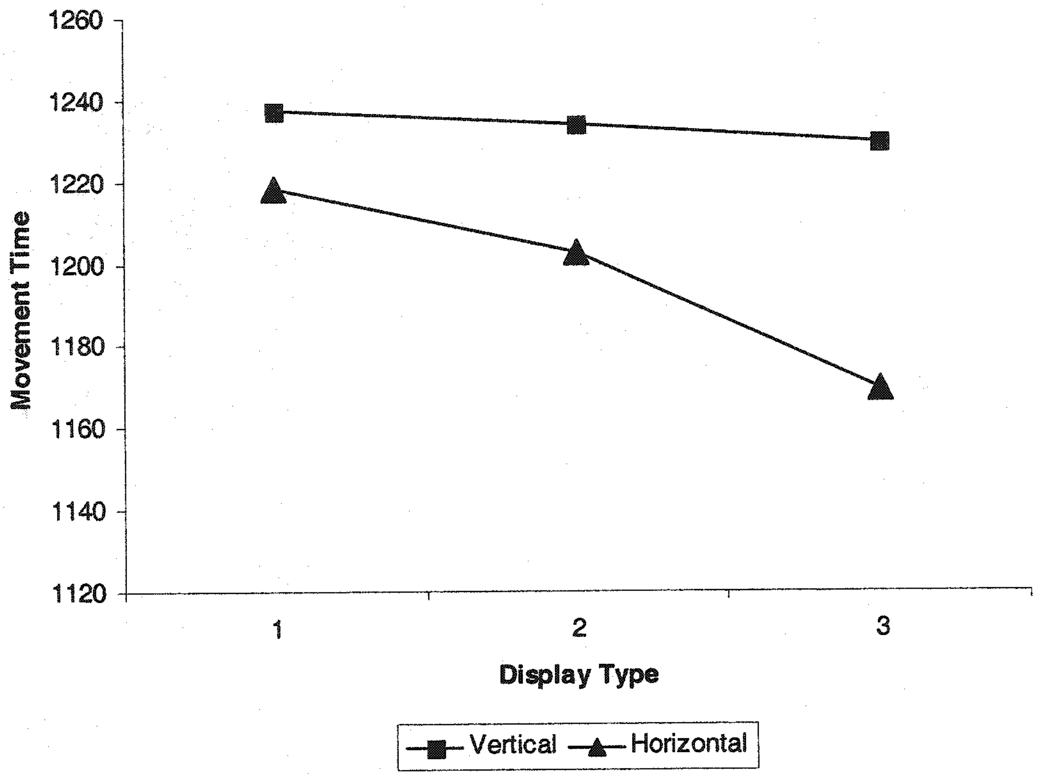
Figure 5. Movement Time for Each Target Size as a Function of Task Difficulty. For each target size, there was a linear relationship between the movement time and the distance.

(e.g. ID = 2.01 vs. ID = 1.74; ID = 3.51 vs. ID = 3.25; ID = 4.25 vs. ID = 4.07; and ID = 5.10 vs. ID = 4.81). With respect to these ID's it is possible that the participants responded faster based on their perception of distance (shorter distance resulted in faster MT), regardless the size of the target object.

With respect to background, the MT was significantly faster for control (M = 1199.74) compared to MT for linear perspective (M = 1227.91) and MT for foreshortening (M = 1218.41) at $p < .05$. The movement times for linear perspective and foreshortening were not significantly different from each other. The Index of Difficulty x Background interaction was not significant ($p > .05$) which suggests that the ID and the background did not interact in their effects on MT. Therefore the MT is affected by the proximal size, not by the distal size of the target. In other words, the observers responded to the proximal properties of the stimuli.

Additional analyses were performed on movement time for each target size using a 3 (background: linear perspective, foreshortening, and control) x 2 (direction of movement: horizontal and vertical) x 3 (distances: 60, 170, and 250 mm) within-subjects ANOVA. This analysis was meant to test and verify the earlier findings that the observers respond to the proximal properties of the stimuli. For each target size there was a significant main effect of distance with the movement time increasing as the distance increased. Therefore, for W = 10mm, $F(2, 54) = 356.59$, $p < .001$; for W = 18mm, $F(2, 54) = 335.29$, $p < .001$; for W = 30mm, $F(2, 54) = 396.52$, $p < .001$; and for W = 36mm, $F(2, 54) = 375.19$, $p < .001$. This confirms the previous findings that the observers respond to the proximal properties of the stimuli. In order to establish how size was perceived at each distance, a 3 (background: linear perspective, foreshortening, and control) x 2 (direction of movement: horizontal and vertical) x 4 (target size: 10, 18, 30, and 36 mm) within-subject ANOVA was conducted on movement time for each distance. At each distance there was a significant main effect of target size (at A = 60mm, *F*(3, 81) = 211.15, $p < .001$, at A = 170mm, *F*(3, 81) = 157.37, $p < .001$, and at A = 250mm, *F*(3, 81) = 246.07, $p < .001$). Post hoc analyses showed that the MT increased as the size of the target decreased regardless of distance.

With respect to the hypothesis that states that the direction of movement (horizontal and vertical) is influenced by the depth cues, the results of the 3 (background: linear perspective, foreshortening, no cue/control) x 2 (two directions of movement: horizontal and vertical) x 12 (ID: the 12 combinations of sizes and distances) within-subjects ANOVA revealed an interaction of Background x Movement Direction (*F*(2, 54) = 3.69, $p < .05$) (see Figure 6). The post hoc analyses showed that for the horizontal approach the MT was significantly faster for control (*M* = 1169.59) compared to MT for linear perspective (*M* = 1218.38) and foreshortening (*M* = 1202.57). For the vertical approach MT was significantly faster for control (*M* = 1229.89) compared to linear perspective (*M* = 1234.25) and foreshortening (*M* = 1237.43) but there was no significant difference of MT for linear perspective and foreshortening. For each background the MT was faster for the horizontal approach compared to the MT for the vertical approach. Thus, our results replicate the earlier findings (Card, English, & Burr, 1978; MacKenzie & Buxton, 1992; Whisenand & Emurian, 1996) that the MT for the horizontal direction is faster than the MT for the vertical direction. We also found that the MT was faster for the



Display Type: 1 = Linear Perspective; 2 = Foreshortening; 3 = Control

Figure 6. The Background x Movement Direction Interaction. For each display type the MT for the horizontal approach was significantly faster than the MT for the vertical approach. For the horizontal approach the MT was faster for the control condition, followed by foreshortening and linear perspective.

horizontal direction even when the depth cues were present. We expected that the MT would be faster in the presence of the depth cues if the observers would respond to the distal properties of the stimuli. The interaction suggests that the depth cues did not influence MT in the direction expected; hence the observers were responding to the proximal properties of the stimuli.

Discussion

The main purpose of the study was to investigate whether there was an impact of the depth cues on movement time as described by Fitts' law. The results showed that the movement time increased as the size of the target decreased and distance to the target increased as predicted by Fitts' law. Since the relationship between the MT and ID was not absolutely linear (see Figure 3) although the trend existed, we could conclude that the distance had a greater effect on MT than the ID (see Figure 5). Since previous research (Epps, 1996) has shown that the trackball (the device used in the study) was the best device found to fit Fitts' law, we could explain that the relatively non-linear relationship may be caused also by the display type. That is, the type of display used in the experiment may not provide enough depth and this may be caused by the absence of the horizon in the visual field. Hence, other formulations of Fitts' law may be used to in order to decide which formula can be applied to best describe the present data. However, the linear relationship for each size was stronger than the overall fit which confirms the idea that the distance had an impact on movement time more than the IDs themselves did.

We also explored whether the depth cues have an impact on how the observers perceive the size, that is, proximally or distally. Target size is included in studies of

Fitts' law applications because it is a part in the equation for the index of difficulty (as explained in the introduction section). Rationally, smaller targets require more exact movements and should therefore take longer to select than larger targets, which is part of what we found in the present study. We examined how this part of Fitts' law equation was affected by the depth cues presented on the display (linear perspective and foreshortening). Moreover, we wanted to examine whether observers perceived the proximal or the distal properties of the stimuli and how this was influenced by the depth cues. We found that regardless of the depth cues, the MT was affected by the proximal size not by the distal size of the target object. That is, the observers relied on the physical size of the target when they made their movements despite the presence of the depth cues.

In vision, the distal stimulus serves as the basis for stimulus energy that travels to the observer and that becomes the proximal stimulus when it reaches the retina of the observer. In other words the 3-D distal stimulus is transcribed into a 2-D proximal stimulus, which brings about a mental representation of the object in 3-D. If the depth cues provided sufficient depth to create a 3-D perspective display, it could have been possible that the observers perceived the distal properties of the object. Therefore, one explanation is that in the present study the depth cues presented on the display did not provide enough depth necessary for the observers to perceive the size distally. However, more research on size judgments is needed in order to determine if the depth cues presented (especially foreshortening) produce size constancy as found by Raddatz, Uhlarik, and Jordan (2001).

A question that arises here is whether the observers perceived the distal properties of the stimuli but they were not able to make the movements associated with them. The theory of the visual perception of objects (ventral stream - projecting from the primary visual cortex to the inferotemporal cortex) and the visual control of action directed at those objects (dorsal stream - projecting from the primary visual cortex to the posterior parietal cortex) relies on different circuits in the human brain has been generally accepted. The theory states that the information about object characteristics (e.g. size, shape, orientation) proceeds through the ventral pathway whereas the information involving egocentric location of objects used for visually guided action proceeds through the dorsal pathway (Goodale & Milner, 1992). Lee and Donkelaar (2002) conducted an experiment in which they disrupted either the ventral or the dorsal stream processing (using transcranial magnetic stimulation - TMS) while subjects made pointing movements as quickly and accurately as possible. The subjects pointed to the central target circles within the Ebbinghaus illusion display and to target circles of physically different sizes that were not surrounded by the either large or small circles (control experiment). With respect to Ebbinghaus illusion, the researchers found that the TMS (over either the ventral or dorsal stream) reduced the influence of illusion on the pointing movement speed but did not affect reaction time or movement accuracy. With respect to the control experiment, they found that the effect on pointing movement was still present with dorsal but not ventral stream stimulation. Their results brought evidence that the ventral stream contributes to pointing movements based on relative object size information via its projection to the prefrontal areas and not through interactions with the dorsal stream. In light of this theory, it is possible that in our study the observers responded to the distal properties of the stimuli but they made their movements based on the proximal properties of the stimuli, since there are separate pathways for perception and action.

Previous research (Van Breda & Veltman, 1998) has shown that the time to perform a target acquisition task was faster when the cue of linear perspective was present. Our findings do not support these reports and we assume that the reason is due to the different depth cues and the different displays used. In their study, Van Breda and Veltman (1998) used a grid surface as perspective which is a combination of linear perspective and foreshortening and the display had a horizon present on the screen. In our study, the perspective was created by either linear perspective or foreshortening and the display did not have a horizon present on the screen.

We also explored whether the direction of movement (horizontal – from left to right and vertical- from down to up) had an impact on movement time in the presence of the depth cues. The MT was always faster for the horizontal approach compared to the vertical approach regardless of the depth cues presence. These results were consistent with those of MacKenzie and Buxton (1992) and Whisenand and Emurian (1996) who did find a significant difference between the MT for horizontal approach and MT for vertical approach.

With respect to the background displays we found that the MT was always faster for control followed by foreshortening and linear perspective and this suggests that the depth cues did not influence the movement time in the expected direction. It is possible that if an effect of depth on the perceived size existed, the MT for the vertical direction might have been about the same as the MT for the horizontal direction. There was no difference between linear perspective and foreshortening.

The trackball was not tested for the horizontal and vertical approaches in past research and therefore, our study brings forth new findings for the trackball. Although it was not our purpose to study the MT for the horizontal and vertical directions as a function of the device, we could expand our findings as follows. When using a trackball, the MT is faster when the target is approached horizontally than MT for the vertical approach, regardless of the presence of the depth cues. This may strengthen the previous research that compared horizontal and vertical MT's across different devices (MacKenzie & Buxton, 1992; Whisenand & Emurian, 1999; Jagacinski & Monk, 1985) and found that the MT was always faster for the horizontal approach compared to the MT for the vertical approach.

In terms of future research we need to examine the perceived size and depth when the depth cues presented on the display have a horizon present on the screen. That is, the horizon will be viewed by the observer, which was not the case of the present study. We expect that the horizon present in the visual field will influence the way observers will perceive depth and hence the size of the objects. Research on size judgments is needed in order to determine whether the depth cues presented in our experiment support size constancy independent of movement of the cursor to a target. Raddatz, Uhlarik, and Jordan (2001), using a different display design, found that foreshortening supports size constancy. Therefore, we need to expand our research to determine whether our displays support size constancy, and whether this can be generalized to other kinds of display designs.

References

Card, S.K., English, W.K., & Burr, B.J. (1978). Evaluation of the mouse, rate-controlled isometric joystick, step keys, and text keys for text selection on a CRT. *Ergonomics, 21*, 601-613.

Epps, B. W. (1986). Comparison of six cursor devices based on Fitts' law models. *Proceedings of the Human Factors Society, USA, 30,* 327-331.

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.

Fitts, P. M. & Peterson, J. R. (1964). Information capacity of discrete motor responses. *Journal of Experimental Psychology, 67*, 103-112.

Gillan, D.J., Holden, K., Adam, S., Rudisill, M., & Magee, L. (1992). How should Fitts' Law be applied to human-computer interaction? *Interacting with Computers, 4*, 291-313.

Goodale, M.A., Milner, A.D. (1992). Separate visual pathways for perception and action. *Trends in Neuroscience, 15,* 20-25.

Graham, E. D. & MacKenzie, C. L. (1996). Physical versus virtual pointing. *Proceedings of the CHI, Conference on Human Factors in Computing Systems,* New York: ACM, 292-299.

Hendrix, C. & Barfield, W. (1995). Relationship between monocular and binocular depth cues for judgment of spatial information and spatial instrument design. *Displays, 16*, 103-113.

Jagacinski, R. J. & Monk, D. L. (1985). Fitts' law in two-dimensions with hand and head movements. *Journal of Motor Behavior, 17,* 77-95.

Jagacinski, R. J., Repperger, D. W., Moran, M. S., Ward, S. L., & Glass, B. (1980). Fitts' law and the microstructure of rapid discrete movements. *Journal of Experimental Psychology: Human Perception and Performance, 6*, 309-320.

Kantowitz, B.H. & Elvers, G.C. (1988). Fitts' law with an isometric controller: Effects of order control and control-display gain. *Journal of Motor Behavior, 20*, 53-66.

Lee, J-H. & Van Donkelaar, P. (2002). Dorsal and ventral visual stream contributions to perception-action interaction during pointing. *Experimental Brain Research, 143,* 440-446.

Luria, S.M., Neri, D.F., & Schlichting, C. (1989). Performance and preference with various VDT phosphors. *Applied Ergonomics, 20*, 33-38.

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

MacKenzie, I. S. & Buxton, W. (1992). Extending Fitts' law to two-dimensional tasks. *Proceedings of the CHI, Conference on Human Factors in Computing Systems,* New York: ACM, 219-226.

MacKenzie, K S. I., Kauppinen, T., & Silfverberg, M. (2001). Accuracy measures for evaluating computer pointing devices. *Computer Human Interaction, 3*, 9-15.

Miller, R.J. (1997). Pictorial depth cue orientation influences the magnitude of perceived depth. *Visual Arts Research, 23*, 97-124.

Murata, A. (1999). Extending effective target width in Fitts' law to a two-dimensional pointing task. *International Journal of Human-Computer Interaction, 11*, 137-152.

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

Raddatz, K., Uhlarik, J., & Jordan, K. (2001). Perceived size in virtual environments: The role of pictorial depth cues. *Proceedings of the Human Factors and Ergonomics Society, USA, 45,* 1404-1408.

Sheikh, I.H. & Hoffmann, E.R. (1994). Effect of target shape on movement time in a Fitts task. *Ergonomics, 37*, 1533-1547.

Van Breda, L. & Veltman, H.A. (1998). Perspective information in the cockpit as a target acquisition aid. *Journal of Experimental Psychology - Applied, 4*, 55-68.

Whisenand, T.G. & Emurian, H.H. (1996). Effects of angle of approach on cursor movement with a mouse: Consideration of Fitts' law. *Computers in Human Behavior, 12*, 481-495.

Whisenand, T.G. & Emurian, H.H. (1999). Analysis of cursor movements with a mouse. *Computers in Human Behavior, 15*, 85-103.

Appendixes

Appendix A. Human Subjects-Institutional Review Board Approval

San José State
UNIVERSITY

**Office of the Academic
Vice President**

***Academic Vice President
Graduate Studies and Research***

One Washington Square
San José, CA 95192-0025
Voice: 408-283-7500
Fax: 408-924-2477
E-mail: gradstudies@sjsu.edu
http://www.sjsu.edu

To: Lucia Arsintescu
335 Elan Village Lane, #316
San Jose, CA 95134

From: Nabil Ibrahim,
AVP, Graduate Studies & Research

Date: April 2, 2003

The Human Subjects-Institutional Review Board has approved your request to use human subjects in the study entitled:

"Effects of Pictorial Depth Cues on Size Perception
in a Target Acquisition Task."

This approval is contingent upon the subjects participating in your research project being appropriately protected from risk. This includes the protection of the anonymity of the subjects' identity when they participate in your research project, and with regard to any and all data that may be collected from the subjects. The approval includes continued monitoring of your research by the Board to assure that the subjects are being adequately and properly protected from such risks. If at any time a subject becomes injured or complains of injury, you must notify Nabil Ibrahim, Ph.D. immediately. Injury includes but is not limited to bodily harm, psychological trauma, and release of potentially damaging personal information. This approval for the human subjects portion of your project is in effect for one year, and data collection beyond April 2, 2004 requires an extension request.

Please also be advised that all subjects need to be fully informed and aware that their participation in your research project is voluntary, and that he or she may withdraw from the project at any time. Further, a subject's participation, refusal to participate, or withdrawal will not affect any services that the subject is receiving or will receive at the institution in which the research is being conducted.

If you have any questions, please contact me at (408) 924-2480.

The California State University:
Chancellor's Office
Bakersfield, Channel Islands, Chico,
Dominguez Hills, Fresno, Fullerton,
Hayward, Humboldt, Long Beach,
Los Angeles, Maritime Academy,
Monterey Bay, Northridge, Pomona,
Sacramento, San Bernardino, San Diego,
San Francisco, San Jose, San Luis Obispo,
San Marcos, Sonoma, Stanislaus

Appendix B. Consent Form

AGREEMENT TO PARTICIPATE IN RESEARCH AT SAN JOSE STATE UNIVERSITY

Responsible Investigators: Lucia Arsintescu, MA in Psychology candidate and Kevin Jordan, Professor of Psychology

I have been invited to participate in research on the perception of visual stimuli in a computer display. There will be a single session to the experiment, lasting less than one hour. I will receive one hour of credit toward my Psychology 001 research requirement based on my participation. Although there my participation in this study will not benefit me directly, I may feel satisfaction that my answers help researchers to understand how the visual system processes size information in a visual display. The possible risk is minimal eye strain equivalent to what might occur from 40 minutes work on a computer. I understand that my participation in the experiment is voluntary.

If I decide to participate, I will be asked to use a trackball to move the cursor into the boundaries of a target specified on the computer display. There will be 300 hundred of these movements for me to make. However, the experiment is quite fast-paced and the entire procedure will take less than 50 minutes.

Data gathered from this study will be stored on a computer disk which no one but the experimenter will be able to access. In case the results of this study are published, any information that is obtained from me in connection with this study and that can be identified with me will remain confidential.

My decision to participate or not participate will not in any way prejudice my future relations with San Jose State University. If I decide to participate, I am free to withdraw my consent and to discontinue my participation at any time without penalty and without any loss of the one hour credit toward my Psychology 001 research requirement for participation in this experiment. I may quit the experiment at any time. No service of any kind, to which I am otherwise entitled, will be lost or jeopardized if I choose not to participate.

If I have any questions, I may ask them prior to the start of the experiment. If I have any questions after the experiment, I may contact Dr. Jordan at 924-5626 or drop by DMH 342. If I have any complaints about the procedure, I may contact Dr. Robert Pellegrini, Psychology department Chair at 924-5600 (DMH 157). For questions about research participants' rights, or in the event of research-related injury, I may contact Dr. Nabil Ibrahim (Associate Academic Vice President for Graduate Studies) at 924-2480.

I am making a decision whether or not to participate. My signature indicates that I have decided to participate having read the information provided above. I have received a copy of this consent form for my records.

SIGNATURE________________________________DATE_____________

PRINTED NAME______________________________________

SOCIAL SECURITY___________________________________

SIGNATURE____
(Investigator)

Appendix C. Instructions

Instructions

Welcome to our experiment. Your task in the experiment is to move a cursor which is in the shape of a triangle into a target which is in the shape of a circle as quickly and accurately as possible. The cursor will be placed either on the left of the screen or on the bottom. In order to complete this task you will use a trackball. To begin the experiment and then to proceed from a block of trials to another you have to press any key on the keyboard. After you pressed the key, the cursor and the object will appear on the screen. Then you have to move the cursor into the target as quickly and accurately as possible. When the cursor will reach the target you will hear a beep and a new trial will begin.

The experiment will involve 288 experimental trials (4 blocks of 72 trials) and 12 practice trials. Before starting the experiment you will proceed through the practice trials in order to familiarize yourself with the task. Then I will answer you any questions you may have about the task requirement. The experiment should take less than 40 minutes. Thank you for participating in the experiment.

Please keep in mind that your task is always to move the cursor into the target as quickly and accurately as possible.

Appendix D. Debriefing

Debriefing

The ability to judge an object in terms of size with satisfactory degree of accuracy is a very important aspect of spatial performance in virtual environments and it depends to a large degree on how these environments can be created to take a more three-dimensional form. An important factor related to this is the motor process involved; i.e., how people interact with a movement device (e.g. a mouse, trackball, isometric joystick) in order to quickly and accurately point toward an object. Fitts' law is a predictor of target acquisition times on movement tasks where the aim is to reach a target as quickly and accurately as possible as a function of target size and target distance. Researchers have found that objects that are smaller and objects that are at a further distance are hard to capture while objects that are bigger and objects that are at a closer distance are easy to capture.

The purpose of the present study is to examine the perceptual motor processes by examining the movement time and the perceived size of the target when different depth cues (such as linear perspective and foreshortening) are present. Linear perspective is a pictorial cue based on the convergence of vertical parallel lines when receding in depth. Foreshortening is a depth cue based on numerous horizontal parallel lines receding in depth.

Therefore, using Fitts' law we want to examine whether perceived size and hence movement time are influenced by the depth cues depicted on the screen.