Pointing at 3D Target Projections with One-Eyed and Stereo Cursors

Robert J. Teather, Wolfgang Stuerzlinger
York University
Toronto, Ontario, Canada
{rteather, wolfgang}@cse.yorku.ca

ABSTRACT

We present a study of cursors for selecting 2D-projected 3D targets. We compared a stereo- and mono-rendered (one-eyed) cursor using two mouse-based and two remote pointing techniques in a 3D Fitts' law pointing experiment. The first experiment used targets at fixed depths. Results indicate that one-eyed cursors only improve screen-plane pointing techniques, and that constant target depth does not influence pointing throughput. A second experiment included pointing between targets at varying depths and used only "screen-plane" pointing techniques. Our results suggest that in the absence of stereo cue conflicts, screen-space projections of Fitts' law parameters (target size and distance) yield constant throughput despite target depth differences and produce better models of performance.

Author Keywords
3D pointing; cursors; selection; Fitts' law.

ACM Classification Keywords
H.5.1 [Information Interfaces and Presentation]: Multimedia Information Systems – virtual reality. H.5.2 [Information Interfaces and Presentation]: User Interfaces – input devices, interaction styles.

INTRODUCTION

Pointing at three-dimensional objects to select them is a fundamental task in 3D user interfaces and is analogous to 2D pointing in graphical user interfaces. However, 3D selection is complicated by a number of issues not found in 2D systems. First, 3D graphics systems use perspective; much like reality, far objects appear smaller, which may influence pointing task difficulty. Second, many graphics systems, including some games and virtual reality, use stereo display. This introduces stereo cue conflicts, and issues such as cursor diplopia (double-vision). Third, there is no universally accepted 3D pointing technique or device, unlike the 2D domain where the mouse is commonly used.

We investigate the interplay between pointing device, technique, and stereo cursors when selecting perspective-scaled 3D targets. Since a primary goal of our research is to develop better methods to directly compare 2D and 3D pointing, we include both mouse and remote pointing in our studies. Both devices are used with both a screen-plane (2D) pointing technique and a depth-cursor (3D) pointing technique. Although the mouse works with many desktop 3D systems [18], it is impractical in immersive VR systems, and 3D trackers are still frequently used (see e.g., [5, 10, 21]). We "bridge the gap" between these types of systems by comparing both classes of device.

A mouse cursor with ray-casting affords selection of 3D objects via their screen-space projections. However, projections of far objects are smaller due to perspective, and such objects may be harder to hit with the mouse or remote pointing. Therefore, we consider the effect of perspective due to target depth. In our first experiment, target depth is constant between targets. Our second study uses varying target depth. We propose to model the effect of perspective with extensions to the 2D formulation of Fitts' law [3] and the ISO 9241-9 standard [7], rather than extending these toward 3D models. We argue that this is more appropriate in such "2.5D" or projected pointing tasks.

Extending our previous work [20], we investigate stereo cursors, primarily at which the depth the *cursor* should be displayed. Simply displaying a stereo cursor in the screen plane yields stereo cue conflicts and cause diplopia when trying to select objects at different depths. In contrast, a one-eyed (mono) cursor, first suggested by Ware *et al.* [23], eliminates stereo cue conflicts by displaying the cursor only to the dominant eye. It is thus also immune to diplopia.

Our contributions are:

- A comparison of one-eyed and stereo cursors, extending Ware [23] with a more robust experimental paradigm. We show that one-eyed cursors improve performance with screen cursors, but hinder ray-based techniques.
- A novel screen-plane ray technique that outperforms standard ray-casting and may be more adaptable to immersive VR/AR systems than mouse pointing.
- Evidence that 2D projected Fitts' law parameters are more appropriate than 3D extensions when using screen-plane techniques.
- Evidence that consistent target depth does not affect performance with screen-plane cursors

RELATED WORK

Ray-based pointing techniques work with either 2DOF (degrees-of-freedom) input devices, like the mouse, or

CHI 2013, April 27–May 2, 2013, Paris, France.

3/6DOF ones, such as 3D trackers. There is much interest in these techniques in both 2D [8, 14, 22] and 3D [5, 9, 10, 17, 24] user interface research. Ray-based techniques test a virtual line originating at the input device for object intersections. Ray techniques are often used with large displays [8] and collaborative systems [14]. A drawback of ray techniques is the relative difficulty in selecting remote objects [9]. Far objects take up proportionally less screen space due to perspective. However, in a static scene, far targets also appear closer together. Thus, according to Fitts' law [3], pointing at objects at the same visual depth from the viewer projected onto a screen should be unaffected by object depth, since both width and distance parameters scale by the same factor. This does not hold if targets are presented at different depths.

Ware and Lowther [23] report that a "one-eyed" cursor outperforms a stereo cursor in 3D selection tasks with a 3DOF tracker. The stereo cursor required matching the position in all three dimensions. Their one-eyed cursor ignored tracker depth and moved the mono-rendered cursor in the screen plane, effectively pointing at object screen projections. Thus, there are large differences between these two techniques and their study did not account for differences in degrees-of-freedom or input device. Our study expands on this by comparing cursor rendering style across both 2 and 3DOF techniques.

Jota et al. [8] investigated eye and device-centric rays and found that device-centric rays are better for pointing, while eye rays are better for tracing/steering. They used only targets at the screen plane. In contrast, we use targets displayed at varying depths in stereo. Argelaguet et al. [1] also looked at similar ray techniques. Their RCE technique maps tracker rotation to the orientation of a ray originating at the user's eye. RCE was significantly faster than standard ray-casting, especially for (partially) occluded targets.

Fitts' Law and Pointing

Fitts' law [3] is an empirical model of the tradeoff between speed and accuracy in pointing tasks. The model is $MT = a+b\times\log_2(D/W+1)$. MT is movement time, D is target distance, and W is target size, while a and b are empirically derived via linear regression. Fitts' law implies that the smaller and farther a target, the more difficult it is to hit it accurately. The log term is known as the index of difficulty (*ID*) and indicates overall pointing task difficulty. The law is associated with a measure of performance known as throughput. Two variations on this are commonly used: one given as $TP = 1/b$, and the other given as $TP = ID_{avg}/MT_{avg}$. A discussion of the merits of both options is omitted for space reasons, but can be found elsewhere [16, 25].

An international standard [7] recommends a post-experiment adjustment for accuracy to fix the error rate to 4% by re-sizing targets to their "effective" size (W_e). This "normalized" accuracy enables computation of (effective) throughput, a measure that incorporates both speed and accuracy. Here $TP = \log_2(D_e/W_e+1)/MT,$ where D_e is the average of the measured movement distances. The effective width, W_e, is computed by projecting the cursor onto the task axis (the line between subsequent targets) and multiplying the standard deviation of the distances by 4.1333. We previously suggested [19] using the distance between the selection ray and the target as a more accurate representation of the effective width W_e, as the actual 3D cursor position may be arbitrarily far away on near misses.

The main advantage of effective measures is that throughput variability for the same condition tends to be lower [16]. Consequently, results of pointing studies are more consistent and comparable. This helps account for the speed-accuracy tradeoff, i.e., optimizing for speed typically hurts accuracy, and vice versa. Throughput scores are largely insensitive to this tradeoff [12].

Although Fitts' law was developed originally for one-dimensional motions, it works extremely well for 2D motions and is commonly employed in the evaluation of pointing device performance [11]. Straightforward extensions to 3D pointing generally focus on improving the correlation between *MT* and *ID*. Note however that adding *any* extra free parameter in a regression analysis will always improve the correlation [15]. Thus, it is not always clear if extra factors improve models' predictive capabilities appropriately.

For example, Murata and Iwase [13] used a 3D tracked device to evaluate pointing tasks on a vertically oriented 2D plane. This was not a true 3D task, as it did not involve hitting targets at varying depths. They derived a model for *ID* incorporating the angle to the target, and report a higher correlation between *MT* and their *ID* model. Grossman *et al.* [4] investigated pointing at tri-variate 3D targets, i.e., targets varying in height, width and depth. Their model considered the direction of movement as a vector through the target. Yet, they used only targets positioned on a single "ground plane" parallel to the floor, effectively a 2D task. They experimentally validated their model using a volumetric display and a tracked input device. Volumetric displays provide more complete depth cues than stereo systems. Moreover, they used a 3D cursor, rather than screen-plane cursors as in our study.

The model presented by Kopper et al. [9] favors the use of "angular" width and distance parameters to model remote pointing. Effectively, this results in targets closer to the user being easier to hit (effectively larger), due to increased ray precision near the ray origin. While the model was validated for 2D target selection tasks on large screen displays, it theoretically will also work for ray-based selection in virtual environments. Our proposed model is similar, in that targets are effectively resized based on the degree of perspective scaling they are subject to.

POINTING TECHNIQUES

While our current work is largely focused on cursor properties, we also consider the effect of the input device, as the two are not independent. Thus, we use two different cursor modes with each device. The first uses a screen plane cursor and the second a sliding cursor [19]. Our first study included all four combinations depicted in Figure 1.

The first mouse technique, which we refer to as MC (mouse cursor), (Figure 1a) displays a cursor in the screen plane and uses the eye-cursor ray for selection. This represents typical 3D selection techniques with the mouse. The sliding mouse cursor, or MS (mouse slide), (Figure 1b) instead displays the cursor where the (same) selection ray intersects the *scene*. Thus, the cursor slides across the geometry. Our novel "ray-screen" technique, RS, (Figure 1c) displays a screen cursor where the device ray intersects the screen, but does not use this ray for selection. Instead, the ray from the eye through this cursor is used for selection. This effectively affords selection of object projections via a user-controlled cursor on the screen, similar to mouse pointing. This is different from Argelaguet et al. [1], who used solely input device rotation to control the cursor. While RS is somewhat similar to zoomable interfaces, it also affords off-axis pointing and uses an implicit zoom control (as a function of perspective scaling). The final technique, RC, (Figure 1d) is traditional ray casting: a device-centric ray that requires users to point the device directly at the 3D target position, which is a form of sliding cursor.

SCREEN PLANE POINTING & PERSPECTIVE SCALING

We hypothesize that selecting targets presented at the same depth yields constant performance. Here *ID*, which depends on *D* and W, is unaffected by target depth as both parameters are scaled by the same factor due to perspective. Perspective scaling of targets is depicted in Figure 2. The same scaling applies to distances as well. Consequently, we expect that screen-plane pointing techniques, such as mouse and ray-screen, are not affected by target depth, assuming a one-eyed cursor is used to avoid diplopia. This does not apply to targets presented at different depths nor when head-tracking is used since both affect how targets project to the screen. In a head-tracked system, ID would constantly change for screen-plane conditions as each head motion could affect the target size and distance. In both cases, the *D* and *W* parameters scale by different factors, and *ID* will subsequently change from what was presented. For small head movements or targets that are far away this change may be insignificant, though.

METHODOLOGY

Here we describe the two user studies we performed to investigate the effect of cursors, devices, and target depth on performance. The first study looks only at cursors and devices for motions between targets at equal depths. The second study investigates a subset of the conditions on motions between targets at different depths.

USER STUDY 1

This study establishes a baseline for essentially 2D target selection of 3D target projections. Consequently, targets were presented at a consistent depth, i.e., in each circle of 11 targets, all targets were at the same visual depth from the viewer. All trials in a circle used the same depth. However, depth was varied *between circles* to determine if performance was constant, despite target depth.

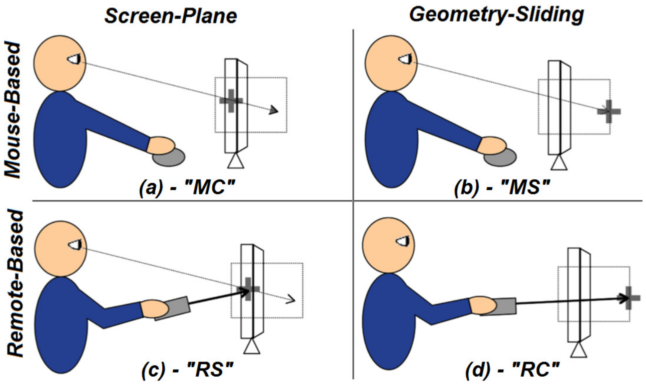

Figure 1. The four 3D pointing techniques used in the first study. Our second study used only the screen plane techniques on the left of the figure: mouse cursor and ray-screen.

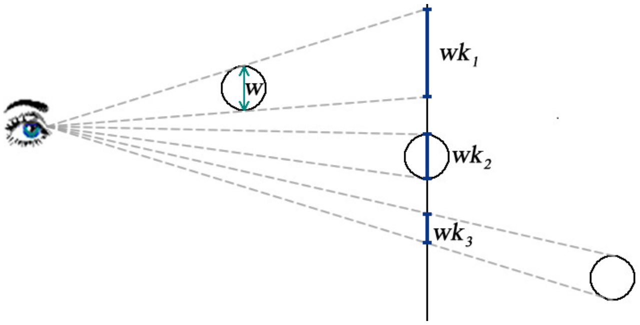

Figure 2. Three targets (circles) positioned at three distinct depths relative to the eye are all of size *w*. They project to different sizes (wk_1, wk_2, and wk_3), nonlinearly depending on distance, onto the display surface, depicted as the black line. The same argument applies to target distance.

Participants

We recruited sixteen paid participants (mean age 23.1 years, *SD* 6.1), all undergraduate students at our university. Eight were female. All use the mouse with their right hand and have normal stereo viewing capability. Six participants had previously used 3D input devices in pointing studies.

Apparatus

We used a 3 GHz PC with Windows XP, an Nvidia *Quadro 4400*, and a 24" 120 Hz stereo LCD. The participant sat approximately 65 cm away from the display on a fixed chair. Although the system supports head-tracking, this was disabled to avoid the potential confounds discussed above. Instead, the user sat in a fixed chair. The stereo LCD was synchronized via an RF hub with NVidia *3DVision Pro* shutter glasses. Five NaturalPoint *Optitrack S250e* cameras were used for 3D tracking. The tracked remote pointing device was calibrated to 0.7 mm RMS. End-to-end system latency was about 65 ms. No smoothing was used, as noise was already very low and the latency cost of additional

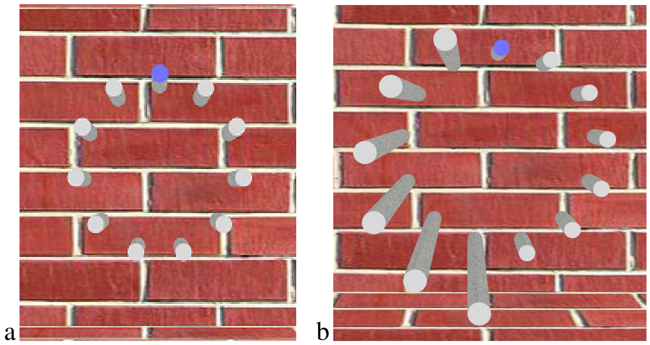

Figure 3 (a). User Study 1 setup: the scene depicting a target circle at –20 cm depth. (b) User Study 2 setup: The scene depicting targets at varying depths. Targets on the right side of the circle are at a depth of –20 cm, while targets on the right are presented at a depth of +8 cm relative to the screen.

filtering may outweigh the benefits [18]. Mouse acceleration was disabled, and gain was set to one level lower than default, for a constant gain of 0.75 [2]. Although low gain levels may increase clutching and impact performance, we rarely observed this in our study.

The 3D scene was a 30 cm deep box matching the display size, see Figure 3a. We used textures and cylinders to facilitate spatial perception of the 3D scene. Target spheres were placed on top of cylinders arranged in a circle. The active target was highlighted in blue. Targets highlighted red when intersected by the cursor. Selection was indicated by pressing a button on the device. The cursor was always displayed as a small 3D crosshair, either at the screen plane or in the 3D scene, depending on the current condition. The center point of the 3D crosshair had to be inside the target sphere for successful selection; otherwise, the software recorded a miss. In one-eyed mode, the cursor was displayed only to the viewer's dominant eye. In ray mode, the 3D device ray was also displayed to improve feedback. Stereo display was active in all conditions, regardless of cursor style. Target size, distance, and depth were constant within target circles, but varied between circles. Target depth was measured relative to the screen surface; negative depth indicates a target behind the screen.

Procedure

Participants were first instructed on the task. To partially compensate for their lack of familiarity with remote pointing, participants were asked to perform 10–20 practice trials with the ray techniques, until they felt comfortable. Participants were instructed to select the blue highlighted target as quickly and accurately as possible. The general experimental paradigm followed that of ISO 9241-9 [7]. Target order started with the "top-most" target (highlighted in Figure 3) and always went across the circle.

Design

This study used a 2×4×4 within-subjects design. The independent variables were cursor style (one-eyed, stereo), technique (MC, MS, RS, RC), and target depth (+8, 0, –8, –20 cm). The dependent variables were movement time (ms), error rate (percentage of targets missed), and throughput (bits per second). There were 10 trials recorded per target circle. Each target circle represented a different index of difficulty, combinations of 3 distances and 2 sizes. Target distances were 7, 15, and 19 cm, while sizes were 0.9 or 1.5 cm. This yields six distinct *IDs* ranging from 2.5 to 4.5 bits, representing a typical range of pointing task difficulty. Each participant completed a total of 1920 trials, for a total of 30720 recorded trials overall.

Results & Discussion

Data were normally distributed according to a Shapiro-Wilks test at the 5% level. Results were analyzed using repeated measures ANOVA and Tukey-Kramer multiple comparisons at the 5% significance level (with Bonferonni correction). Statistical results are reported in Table 1.

Movement time

Overall there was a significant main effect of technique on time. Movement times are shown in Figure 4. Both mouse techniques were significantly faster than the remote pointing ones. Ray-screen was significantly faster than ray-casting. On average, the one-eyed cursor ($\mu = 1321$ ms, $\sigma = 554$ ms) increased movement time compared to the stereo one ($\mu = 1211$ ms, $\sigma = 839$ ms). However, there are strong interaction effects with technique, as ray-casting was far worse with the one-eyed cursor.

Effect		**Movement Time**		**Error Rate**		**Throughput**	
	d.f.	*F*	*p*	*F*	*p*	*F*	*p*
(T)ech	3, 15	62.7	***	13.5	***	103.1	***
(C)ursor	1, 15	16.9	**	3.4	.08	0.26	ns
(D)epth	3, 15	7.4	**	6.1	*	18.1	***
T×C	3, 45	46.7	***	8.7	**	52.0	***
T×D	9, 45	11.7	***	2.0	*	7.7	***
C×D	3, 45	13.4	***	2.9	*	5.3	**
T×C×D	9, 135	4.3	***	0.97	ns	4.9	***

Table 1. User Study 1 statistical report. Significant effects are marked * for *p* <. 05, ** *for p* < .001 and * for *p* < .0001.**

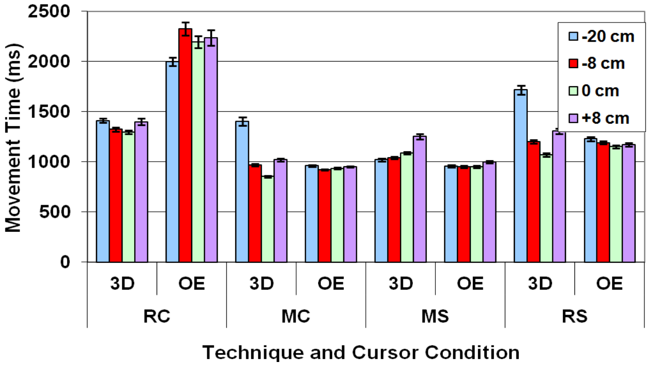

Figure 4. Movement time for each condition. One-eyed cursor conditions are represented with "OE", stereo cursor conditions with "3D". Error bars show ±1 standard error.

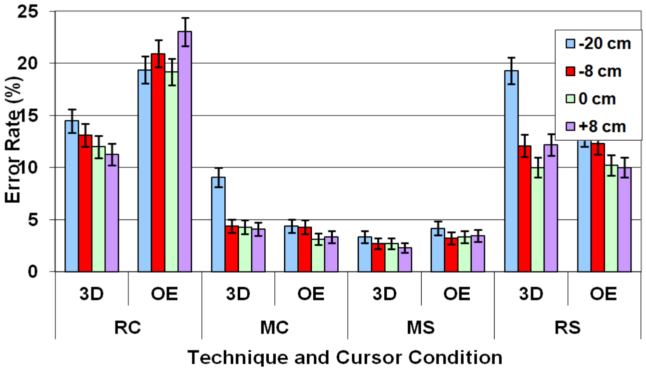

Figure 5. Error rates for each condition. Error bars show ±1 standard error

There was a significant two-way interaction between cursor style and technique. Ray-casting with the one-eyed cursor was significantly worse than all other conditions. The other conditions all *benefitted* from the one-eyed cursor. A significant three-way interaction effect between technique, cursor style, and target depth reveals that the screen-plane conditions (mouse and ray-screen) with the stereo cursor performed significantly worse at the -20 cm depth.

Error Rate

Error rate is the percentage of trials where the participant missed the target. There was a significant main effect of technique on error rate, see Table 1 and Figure 5 for error rates. Both mouse techniques had significantly lower error rates than both remote techniques, around 4%, consistent with 2D pointing experiments. A significant interaction between technique and cursor style revealed that the one-eyed cursor increased error rates with the RC technique.

Throughput

(Effective) throughput was computed as described earlier. There was a significant main effect of technique on throughput, see Table 1 and Figure 6. A Tukey-Kramer test revealed three groups: both mouse conditions were close to 4 bits per second and consistent with 2D pointing, followed by RS at around 3 bps, and finally RC at 2.5 bps. Cursor style alone did not affect throughput. However, there was a significant interaction effect between technique, cursor style, and target depth. Throughput fell dramatically for targets at -20 cm depth with the stereo cursor for both the mouse and ray-screen conditions. The one-eyed cursor hindered the RC technique, which was the worst condition overall, regardless of target depth.

Modeling

Fitts' law can also be used as a predictive model, by regressing movement time on index of difficulty. We performed this analysis for each technique for both the stereo and one-eyed cursor, as presented in Figure 7 and Figure 8. The predictive quality of the model (as expressed by the R^2 values) is very high. However, it is worth noting that the one-eyed cursor consistently improved R^2 values. The one-eyed mouse cursor conditions both show $R^2 \approx 0.97$, indicating almost perfect prediction for both the MC and MS conditions. The stereo cursor degrades the correlation, especially for the MC and RS techniques, likely due to the more pronounced stereo conflicts on deeper targets. The sliding cursor is affected much less, likely because cursor and target depths are the same most of the time. Overall, this illustrates that the predictive capabilities of Fitts' law are unaffected by target depth for 3D pointing techniques that use 2DOF input and a 2D cursor visualization.

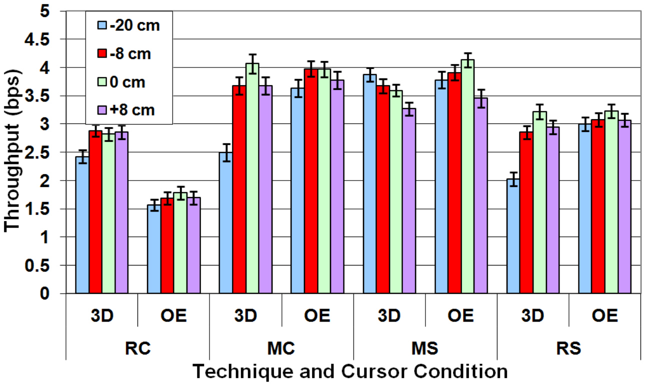

Figure 6. Throughput for each condition. Error bars show ±1 standard error. Higher throughput is better.

Discussion

Consistent with previous results [23], the one-eyed cursor improved performance, but only for certain pointing techniques. Only the mouse, mouse-slide, and ray-screen

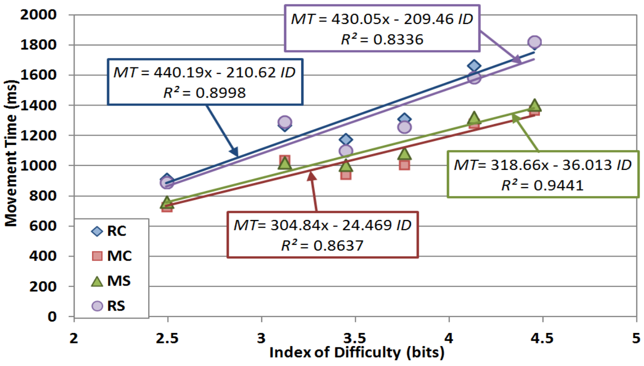

Figure 7. Fitts' law models for stereo cursor conditions.

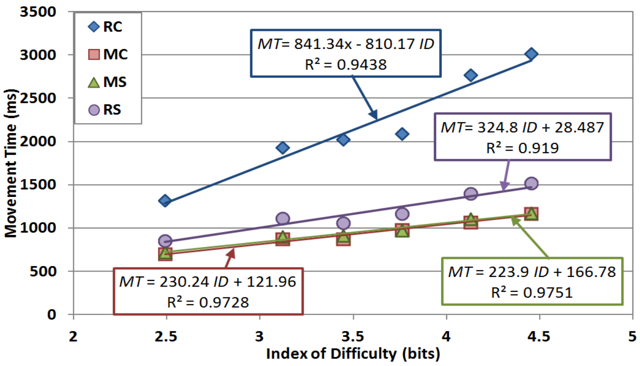

Figure 8. Fitts' law models for one-eyed cursor conditions.

conditions benefitted, while ray-casting performed worse with the one-eyed cursor. Our results also quantify the benefits of the one-eyed cursor in a more robust experimental paradigm compared to the original [23].

The one-eyed cursor improved performance with mouse-based techniques by reducing the impact of target depth in these conditions. The depth effect is most noticeable in the screen-plane stereo cursor conditions. In particular, throughput peaked at 0 cm depth (i.e., at the screen surface) and fell for targets at different depths. The +8 cm and −8 cm depths show similar throughput, but the −20 cm condition shows a dramatic degradation of performance. This is likely due to diplopia. The one-eyed cursor does not suffer from this problem, as it eliminates cursor depth cues altogether, and is thus immune to diplopia. See Figure 6.

Movement time for the mouse slide technique using the stereo cursor was significantly faster for deeper targets compared to closer ones. This seems to be related to participants sliding the cursor up the sides of the target cylinder instead of relying on it “popping” to the front. We previously observed this suboptimal behavior when using sliding cursor techniques [18]. The one-eyed cursor *eliminated* this problem, and participants reported that they could not tell the difference between that condition and the one-eyed mouse (screen) condition. The movement times for these conditions are nearly identical independent of depth, and are not significantly different ($F_{1,15} = 0.23$, ns).

Our results reveal also the differences between pointing techniques. The mouse techniques performed best, but the new ray-screen technique was competitive and significantly outperformed standard ray-casting. We thus recommend this style of image plane technique over classical ray-casting for VR systems and games alike. This is similar to Argelaguet’s results [1], but does not agree with Jota’s work [8]. Our study used a stereo desktop VR system, while Jota used a large non-stereo display system. This difference may account for the discrepancy and our results may thus not generalize to large displays. The multiple interaction effects indicate that most techniques work best with a one-eyed cursor, while some require a stereo cursor. Similarly, some techniques perform best for deeper targets, while others perform best for close targets.

Finally, the one-eyed mouse cursor afforded throughput similar to a standard 2D mouse cursor. This was fairly consistent for both one-eyed mouse conditions. The one-eyed ray-screen condition was also unaffected by target depth. The movement times confirm that performance is unaffected by the perspective scaling of a scene with targets displayed *at the same depth* when using screen-plane techniques. The following study expands our investigation by looking at pointing for targets *at different depths*.

USER STUDY 2

In this study, target depth varied between subsequent targets. As a result, perspective distortion affected the projection of the targets. The objective of this study was to empirically measure and model the effect of perspective. To keep the size of the study manageable, we included only the best-performing mouse and ray techniques from the first study, i.e., the mouse cursor and ray-screen conditions.

Participants

Twelve participants (mean age 29.4 years, *SD* 5.7) took part in the study. Nine were male, and all were right-handed.

Apparatus

The hardware setup was identical to that used in Study 1. However, the software was modified such that target depth varied from target to target. Each target circle was arranged such that every *other* target was at a different depth. This ensured that every subsequent target selection required moving either from a deep target to a near target, or vice versa. This can be seen in Figure 3b. Correspondingly, data were later split into “up” and “down” motions to analyze each separately. This design is one of the few options for accurately analyzing 3D movements with the ISO standard, which requires uninterrupted “circles” of targets.

Procedure

While the apparatus was modified, the procedure was identical to that of the previous study.

Design

The study used a 2×2×3×3 within-subjects design. The first two independent variables were cursor style (one-eyed, stereo) and technique (MC, RS). The remaining independent variables were all nine possible combinations of the three target depths (+8, 0, −20 cm). The dependent variables were movement time (ms), error rate (percentage of targets missed), and throughput (bits per second). There were 12 trials recorded per target circle. Each target circle represented a different index of difficulty, combinations of 3 distances and 2 sizes. Target distances, more precisely the distances between cylinders, were 7, 15, and 19 cm apart, while target sizes were 0.9 or 1.5 cm in diameter. This yielded six distinct *IDs* ranging from 2.5 to 4.5 bits, when computed according to the conventional formulation of Fitts’ law (discussed further below). Thus each participant completed a total of 2592 recorded trials, for a total of 31104 trials overall.

Results & Discussion

Approximately 8% of all trials were dropped as outliers. Trials were considered outliers if their movement times were more than three standard deviations from the grand mean time. After outlier removal, the data were normally distributed according to a Shapiro-Wilks test at the 5% level. Results were analyzed using repeated measures ANOVA and Tukey-Kramer multiple comparisons at the 5% significance level (with Bonferonni correction). We separated the data for each “round” of trials with different target depths into two sets: upwards and downwards movements and treated these two separately from then on, including the calculation of standard deviations.

Effect		Movement Time		Error Rate		Throughput	
	d.f.	*F*	*p*	*F*	*p*	*F*	*p*
(T)ech	1, 11	16.7	*	37.0	***	9.3	*
(C)ursor	1, 11	3.1	.11	8.8	*	1.2	.30
(D)epth	8, 11	16.9	***	17.3	***	9.4	***
T×C	1, 11	1.1	.32	3.1	.11	0.03	ns
T×D	8, 88	3.8	***	6.9	***	1.1	.39
C×D	1, 88	11.7	***	4.2	***	10.3	***
T×C×D	8, 431	1.5	.19	2.5	*	1.8	.08

Table 2. Statistical results for User Study 2. Significant effects are marked * for $p <. 05$, ** *for* $p < .001$ and * for $p < .0001$. Depth represents all combinations of target depths.**

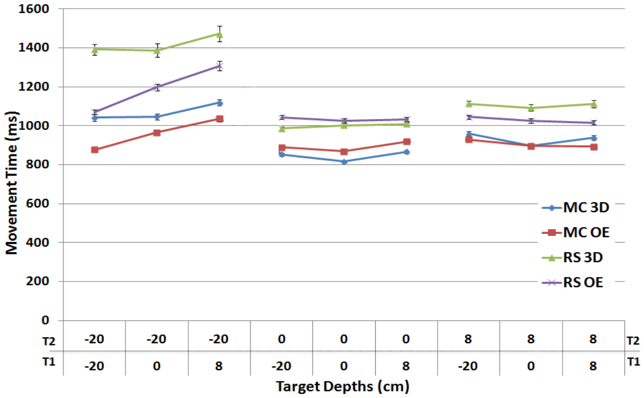

Figure 9. Movement time by depth combination, cursor mode, and input technique for study 2. Error bars show ±1 standard error. T1 and T2 are the start and end depths of each trial.

Movement time

Movement times are shown in Figure 9 and statistical values are shown in Table 2. Technique had a significant main effect on movement time, while cursor style did not. However, depth combination did have a significant effect, suggesting that it was a greater source of variability. Significant interactions between technique and depth suggest that ray-screen is more strongly affected by increasing movement *into* the scene. This is likely because these targets are perspective-scaled to appear smaller while target distance stays (mostly) constant, and the ray technique is subject to greater input device noise. An interaction between depth and cursor style suggests that stereo cursor performance falls with deep targets, regardless if the target depths are the same (e.g., the –20 to –20 condition) or not. The slowest conditions overall were ray-screen with stereo cursor and motions involving –20 cm deep targets. Movement *out* of the scene or in front of the screen had relatively little impact on performance, regardless of technique or cursor style, see Figure 8. The fastest condition was the ray-screen/stereo cursor condition at the screen surface, i.e., when all targets were at 0 cm.

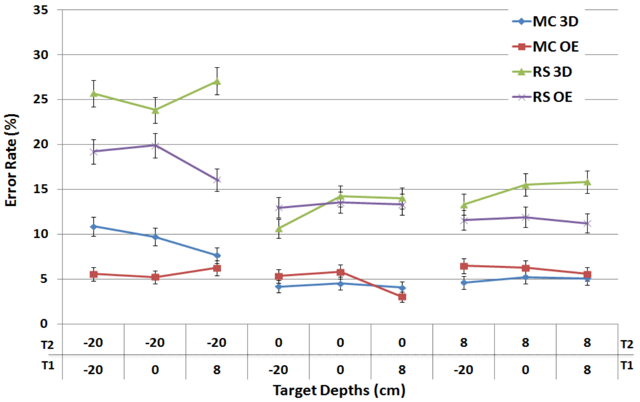

Figure 10. Error rates by technique, cursor style, and depth combination. Error bars show ±1 standard error. T1 and T2 are the start and end depths of each trial.

Error Rate

Error rate is the average percentage of trials where participants missed the target for a given condition. Statistical results for error results can be found in Table 2. Error rates are summarized for each condition in Figure 10. Every investigated condition had a significant main effect on error rate. The error rate for the one-eyed mouse cursor is around 5.5%, slightly higher than in the previous study.

For the RS technique the average error rate is much higher than for the mouse, between 10% and 25%. This is highlighted by the significant interaction effect between the technique, depth, and cursor conditions. While the ray-screen condition is significantly worse than the mouse cursor, it is unsurprisingly far worse with a stereo cursor when pointing at deep targets. This can be seen in Figure 10 for any target depth ending at a –20 cm target. On the other hand, the mouse cursor error rate is essentially constant with the one-eyed cursor, regardless of the depth of the start or end target. This is further evidence that this condition is unaffected by target depth.

"Euclidean" Throughput

Initially, we computed throughput as in the previous study and also in our previous work [19], by using the 3D distances between the target and the closest point of the ray from the eye through the cursor position at the "click". Yet, this artificially inflates throughput scores for movements with greater depth differences. In Figure 11 this manifests as a "dip" in the middle, with inexplicably higher throughput scores for greater depth differences. One can see a similar "dip" in some conditions in Figure 12 in previous work [19], especially for ray-casting.

Screen-projected Throughput

To avoid this inflation, we argue that for pointing techniques that require effectively only 2D input, such as mouse cursor, ray-screen, and to a large degree ray-casting, performance should be evaluated in the screen plane. This motivates the development of a new screen-projected throughput score, which first projects the pointing task onto the screen plane.

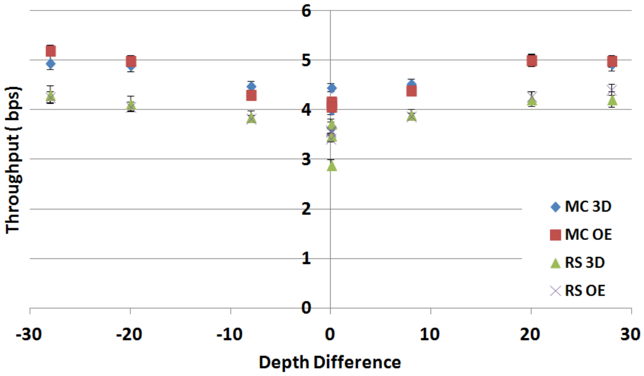

Figure 11. "Euclidean" throughput, illustrating the artificial inflation of throughput scores for movements with greater depth differences.

For this, we first project target and cursor positions to the screen plane, depending on target depth. Effective width is then computed using the standard deviation of the *2D* distances from the projected target to the projected cursor instead of 3D distances. For simplicity, we ignore the (small) effect of perspective onto the shape of the target spheres. Effective distance is computed as the 2D distance between the two projected clicks for each trial. Throughput is then computed normally from these values. The statistical results for screen-projected throughput are shown in Table 2 and mean throughput scores can be found in Figure 12.

Technique had a significant main effect on the new screen-projected throughput while cursor style did not. The combination of start and end target depth did have a significant effect as well. Overall, the mouse cursor affords significantly higher throughput than the ray-screen technique. There is a significant interaction effect between cursor style and target depth combination. Pointing at deeper targets is significantly worse with the stereo cursor than with the one-eyed cursor. The end target depth of the current trial (T2) seems to matter most here. For example, throughput is fairly consistent for all −20 cm deep targets, irrespective of the depth of the start target depth (T1).

Discussion

Screen plane throughput was not affected by depth with either technique (MC OE, or RS OE). To reiterate the results of the Tukey-Kramer post-hoc test, target depth does not significantly affect the mouse ($F_{6,101} = 0.96$, ns), nor ray-screen ($F_{6,101} = .85$, ns). In the absence of diplopia this suggests that perspective scaling of targets due to depth does not affect pointing performance. This makes sense and supports our argument that screen-projected throughput is an appropriate measure for such tasks. Similar to how throughput behaves for changing distances and sizes in 2D and considering the properties of the pointing techniques when used with a one-eyed cursor, throughput *should* remain constant regardless of target depth. Note that head position may affect this. But we did not consider this in our studies, also because participants did not move their heads much during the experiment.

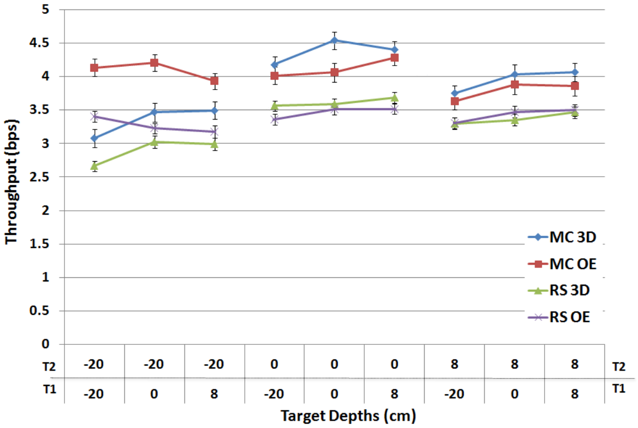

Figure 12. Screen-projected throughput by technique, cursor style, and depth combination. Error bars show ±1 standard error. T1 and T2 are the start and end depths of each trial.

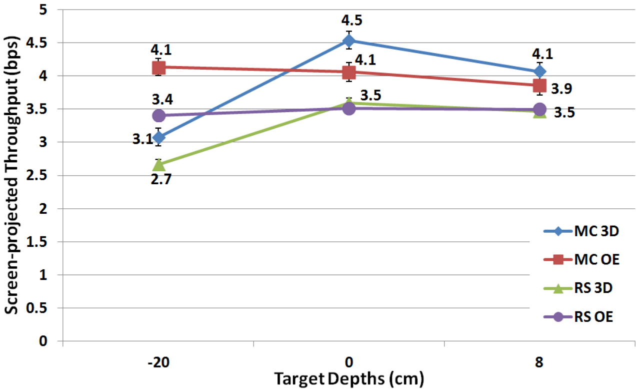

Figure 13. Screen-projected throughput by pointing technique and cursor style for each combination of same-depth targets. Error bars show ±1 standard error.

Perspective Scaling of Same-Depth Targets

We hypothesized that selecting targets subject to the same perspective scaling should yield constant performance when using screen-plane techniques. Hence, if all targets in a circle are at the same depth, then throughput should not change regardless of depth. To verify this, we analyzed this effect in user study 2 for the same-depth conditions, i.e., [−20 to −20], [0 to 0], and [+8 to +8 cm]. Figure 13 depicts the mean screen-projected throughput for each condition.

Figure 13 illustrates that performance for both techniques was mostly constant with the one-eyed cursor. There is at most 5% variation in throughput for the mouse and only 1% for ray-screen. Neither are significant (mouse, $F_{2,33} = 0.3$, ns; ray-screen, $F_{2,33} = 0.16$, ns). While this does not conclusively prove that depth has no effect, it indicates that we cannot reject the null hypothesis – that there is no difference due to depth – and we conclude that this is currently the best explanation for our data. Performance was much more variable with the stereo cursor for both pointing techniques. Due to the stereo cue conflicts present in these cases, this is not unexpected. In particular, the -20 cm depth condition was strongly affected by diplopia, as in our first user study.

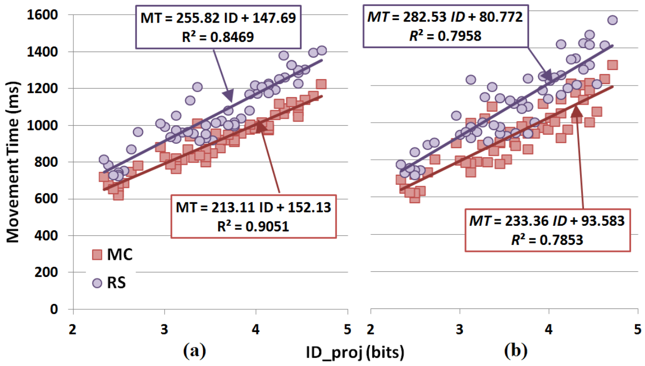

Figure 14. Regression models for the one-eyed (a) and stereo cursor conditions (b) using *ID_proj*, i.e., screen-projected ID. These models include all target depth combinations.

Depth Diff. (cm)	Target Depths (cm)	Mouse, One Eyed Cursor			Ray-Screen, One-Eyed Cursor		
		Intercept (*a*)	Slope (*b*)	R^2	Intercept (*a*)	Slope (*b*)	R^2
-28	8, -20	254.5	198.3	.8554	435.7	199.9	.8113
-20	0, -20	253.5	189.9	.9066	188.2	257.9	.8059
-8	8, 0	77.2	230.8	.9826	52.6	268.5	.9355
0	-20, -20	166.8	198.8	.9551	154.5	252.8	.8781
0	0, 0	111.4	211.3	.9962	119.7	251.0	.9042
0	8, 8	10.3	246.59	.9826	18.4	279.4	.9601
+8	0, 8	54.7	240.3	.9886	-15.4	296.2	.9309
+20	-20, 0	185.0	204.8	.9685	208.5	242.6	.9236
+28	-20, 8	277.9	191.8	.9585	312.7	214.9	.9074

Table 3. Regression models between projected ID and movement time for the one-eyed cursor conditions for each distinct depth difference. Target depths indicate the starting and ending depth of a pointing task.

Modeling

We produced models for each condition using screen-projected target size and distances. We do not incorporate any additional parameters, as screen-projected ID should be sufficient to explain the effect of perspective scaling. Figure 14 presents the aggregate models for each pointing technique, using both one-eyed (Figure 14a) and stereo (Figure 14b) cursor styles.

The models fit slightly worse than one would expect of Fitts' law. This may be due to the time required to re-adjust the eyes to different depths in presence of accommodation-vergence conflicts [6]. Therefore, we performed separate regression analyses for each target depth combination. These models are summarized in Table 3. As expected [6], participants required more time to adjust for greater depth differences. This is visible both as higher intercepts and worse predictive qualities, R^2. The models fit well for near-screen conditions, where depth cue conflicts are weakest.

We did not include regression analyses for the stereo cursor conditions, as the effect of diplopia is too strong to produce reasonable models. Moreover, we are unaware of a model to predict the additional time required to acquire a target in the presence of both diplopia and the aforementioned accommodation-vergence conflict. We intend to investigate this further in future work.

OVERALL DISCUSSION

We did not directly compare our models to others (e.g., [4, 9, 13]) for several reasons. First, our task is essentially 2D, as it involves pointing at 2D projections of targets on the screen plane. Murata's model [13] may be applicable, but we feel their addition of a free parameter is not well justified. There several differences between our task and those used by Grossman [4] or Kopper [9]. In contrast to Grossman's work, we did not use a position-controlled 3D cursor and trivariate targets. Consequently, Kopper's work is a better comparison point. However, a primary objective of our work was a direct comparison of mouse and ray techniques. Kopper's model focused exclusively on remote pointing, and thus likely does not apply to the mouse. Hence, a direct comparison of models between devices is not feasible. Additionally, they used neither stereo display nor varying target depths. Thus their results are not subject to the stereo cue issues we observed. Although we found somewhat lower correlations, our mouse model matches or exceeds their model's predictive capabilities for ray-casting for individual depth conditions.

Implications for Designers

Our results show that 3D user interface designers should be wary of using stereo cursors for selecting targets displayed away from the screen. Interestingly, both studies seem to indicate that stereo cursors offer slightly better performance for targets near the display surface. However, screen-based stereo cursors hurt performance when targets presented away from the screen. This is likely due to diplopia and/or the accommodation-vergence conflict. Our second study suggests that it is the intended target depth that matters most, rather than the actual depth difference.

This also suggests that developers of stereo 3D games should avoid screen-plane stereo cursors. Unfortunately, they are currently common practice in games. Overall, both studies indicate that the advantages of stereo cursors are minimal. But, in general, their usage can significantly *hinder* user performance in 3D pointing. Thus, we recommend that developers consider including a one-eyed cursor option. This leaves the decision of whether to use a stereo cursor to the user, and permits them to avoid performance degradation in stereo display systems.

Finally, there is now interest in the development of stereo touchscreen interfaces [21]. Such interfaces suffer the same problems when interacting with stereo targets far from the screen. Much like a stereo mouse cursor, a finger on a stereo touchscreen is also subject to diplopia! Our work indicates how much of an impact this effect may have.

CONCLUSION

We conducted two studies to investigate stereo cursor properties and the effect of perspective on target selection. Our results quantify the benefits of the one-eyed cursor in a more well-refined experimental paradigm compared to previous work [23] and suggest that the one-eyed cursor is not universally beneficial. We also provide evidence that consistent target depth does not affect pointing performance. Our second study identified that varying target depth affects performance, but this can be (at least partly) accounted for by using screen-plane projections of targets. Overall, mouse-based techniques tended to perform best. But our new "ray-screen" selection technique also outperforms traditional ray-casting! Consequently, we suggest adaptation of this new technique for immersive 3D systems that use remote pointing devices.

Future Work

To investigate the effect of perspective in isolation, we plan to reproduce our study using only mono display. While our current "screen-projected" model fits the data well, is does not account for stereo conflicts. Thus, we are also planning to incorporate a new term into the model to account for the cost of stereo cue conflicts during depth movements.

Acknowledgements

Thanks to the anonymous reviewers for their helpful suggestions to improve the final paper. Thanks to Andriy Pavlovych for proof-reading the final version.

REFERENCES

1. Argelaguet, F. and Andujar, C., Efficient 3D pointing selection in cluttered virtual environments, *IEEE Computer Graphics and Applications, 29*, 2009, 34-43.
2. Casiez, G. and Roussel, N., No more bricolage!: methods and tools to characterize, replicate and compare pointing transfer functions, *ACM UIST 2011*, 603-614.
3. Fitts, P. M., The information capacity of the human motor system in controlling the amplitude of movement, *Journal of Experimental Psychology*, *47*, 1954, 381-391.
4. Grossman, T. and Balakrishnan, R., Pointing at trivariate targets in 3D environments, *ACM CHI 2004*, 447-454.
5. Grossman, T. and Balakrishnan, R., The design and evaluation of selection techniques for 3D volumetric displays, *ACM UIST* 2006, 3-12.
6. Hoffman, D. M., Girshick, A. R., Akeley, K., and Banks, M. S., Vergence-accommodation conflicts hinder visual performance and cause visual fatigue, *Journal of Vision*, *8(3)*, 2008, 1-30.
7. ISO 9241-9, Ergonomic requirements for office work with visual display terminals (VDTs) - Part 9: Requirements for non-keyboard input devices. International Organization for Standardization, 2000.
8. Jota, R., Nacenta, M. A., Jorge, J. A., Carpendale, S., and Greenberg, S., A comparison of ray pointing techniques for very large displays, *Graphics Interface 2010*, 269-276.
9. Kopper, R., Bowman, D. A., Silva, M. G., and McMahan, R. P., A human motor behavior model for distal pointing tasks, *International Journal of Human-Computer Studies*, *68*, 2010, 603-615.
10. Kunert, A., Kulik, A., Lux, C., and Fröhlich, B., Facilitating system control in ray-based interaction tasks, *ACM VRST 2009*, 183-186.
11. MacKenzie, I. S., Fitts' law as a research and design tool in human-computer interaction, *Human-Computer Interaction*, *7*, 1992, 91-139.
12. MacKenzie, I. S. and Isokoski, P., Fitts' throughput and the speed-accuracy tradeoff, *ACM CHI 2008*, 1633-1636.
13. Murata, A. and Iwase, H., Extending Fitts' law to a three-dimensional pointing task, *Human Movement Science*, *20*, 2001, 791-805.
14. Oh, J.-Y. and Stuerzlinger, W., Laser pointers as collaborative pointing devices, *Graphics Interface 2002*, 315-320.
15. Pagano, R. R., *Understanding statistics in the behavioural sciences*. Wadsworth Publishing, 2006.
16. Soukoreff, R. W. and MacKenzie, I. S., Towards a standard for pointing device evaluation, perspectives on 27 years of Fitts' law research in HCI, *International Journal of Human-Computer Studies 61*, 2004, 751-789.
17. Steinicke, F., Ropinski, T., and Hinrichs, K., Object selection in virtual environments using an improved virtual pointer metaphor, *Computer Vision and Graphics*, vol. 32, *Computational Imaging and Vision*, Springer Netherlands, 2006, 320-326.
18. Teather, R. J., Pavlovych, A., Stuerzlinger, W., and MacKenzie, I. S., Effects of tracking technology, latency, and spatial jitter on object movement, *IEEE 3DUI Symposium 2009*, 43-50.
19. Teather, R. J. and Stuerzlinger, W., Pointing at 3D targets in a stereo head-tracked virtual environment, *IEEE 3DUI Symposium 2011*, 87-94.
20. Teather, R. J. and Stuerzlinger, W., Cursors for 3D Pointing, in *3DCHI Workshop*, 2012.
21. Valkov, D., Steinicke, F., Bruder, G., and Hinrichs, K., 2d touching of 3d stereoscopic objects, ACM CHI 2011, 1353-1362.
22. Wang, G., McGuffin, M. J., Berard, F., and Cooperstock, J. R., Pop-up depth views for improving 3D target acquisition, *Graphics Interface 2011*, 41-48.
23. Ware, C. and Lowther, K., Selection using a one-eyed cursor in a fish tank VR environment, *ACM Transactions on Computer-Human Interaction*, *4*, 1997, 309-322.
24. Wingrave, C. A., Tintner, R., Walker, B. N., Bowman, D. A., and Hodges, L. F., Exploring individual differences in raybased selection; strategies and traits, *IEEE VR 2005*, 163-170.
25. Zhai, S., Characterizing computer input with Fitts' law parameters—the information and non-information aspects of pointing, *International Journal of Human-Computer Studies*, *61*, 2004, 791-809.