Factors Affecting Mouse-Based 3D Selection in Desktop VR Systems

Robert J. Teather
Dept. of Computing & Software
McMaster University
Hamilton, ON, Canada
teather@mcmaster.ca

Wolfgang Stuerzlinger
School of Interactive Arts + Technology (SIAT)
Simon Fraser University
Vancouver, BC, Canada
w.s@sfu.ca

ABSTRACT

We present two experiments on mouse-based point selection in a desktop virtual reality system using stereo display and head-tracking. To address potential issues of using a mouse cursor with stereo display, we also evaluate the impact of using a one-eyed (mono) cursor. While a one-eyed cursor visualization eliminates depth conflicts, recent work suggests it offers worse performance than stereo cursors, possibly due to eye discomfort. Our results indicate that presenting the cursor in stereo significantly reduces performance for targets at different depths. The one-eyed cursor eliminates this effect, offering better performance than both screen-plane and geometry-sliding cursors visualized in stereo. However, it also performed slightly worse than stereo cursors in situations *without* depth conflicts. Our study suggests that this difference is not due exclusively to the relative transparency of such a cursor, hence eye fatigue or similar may be responsible.

Categories and Subject Descriptors

H.5.2 [Information Interfaces and Presentation]: User Interfaces – input devices, interaction styles.

Keywords

Mouse cursors, stereo 3D display, head-tracking.

1. INTRODUCTION

Despite falling costs and advances in 3D trackers, the mouse remains the predominant input device for interacting with 3D content. Stereo displays are becoming more common and are effective for visualizing 3D scenes. Low-cost trackers (e.g., Kinect, Wiimote) enable head-tracking, which offers the possibility of head-based viewpoint control, leaving the mouse free for other operations. There is thus interest in using the mouse with so-called "desktop" VR systems – small scale VR systems using stereo display and head-tracking for improved depth cues when interacting with 3D content. Recent work has suggested that there are benefits to stereo (and other) depth cues in otherwise 2D experiences, such as desktop 3D design [4], gaming [9, 15] and general point selection [17]. To this end, researchers have studied best practices in mouse-based input on stereo displays [12, 17, 19]. The potential benefits of head-tracking with mouse-based interaction are comparatively underexplored.

A major problem with using a standard (system) mouse cursor in stereo scenes is that the cursor is presented to both eyes with 0 disparity, yet occludes geometry behind it – even geometry closer to the viewer than the cursor. This yields a mismatch between the perceived cursor depth and that of the occluded geometry. Another problem is double vision (diplopia), caused by aligning the cursor with a feature at a different depth. A simple solution is to display the cursor to only one eye, but recent work suggests this may offer worse performance than 3D cursors [12].

We present two experiments investigating how several factors influence mouse-based selection performance on desktop VR systems. These factors include the presence of stereo display, head-tracking, cursor movement (screen-plane vs. "sliding"), and cursor visualization (stereo, one-eyed, or transparent). The first experiment focuses on depth cues in mouse-based systems. There are two aspects to this experiment. First, we evaluate the effects of common techniques for improving depth perception (e.g., stereo and head-tracking) on mouse-based selection. Second, we examine the one-eyed cursor to quantify potential negative effects previously reported [12]. The main contribution of this experiment is a more systematic investigation of these factors than in previous work [12, 17], which primarily focused on various cursor or interaction styles. The second experiment extends this to assess if differences with a one-eyed cursor may be due to its apparent transparency. Since it is shown to only one eye, it appears effectively 50% transparent compared to a stereo cursor. The study also investigates the relative performance of so-called "sliding" cursors [17] compared to screen-based cursors.

Our experiments fix the screen-based cursor at zero depth/parallax. Note that this choice for the cursor depth is an intentionally sub-optimal one. Previous work [12] shows benefits for matching the parallax of the cursor to the geometry – this motivates our use of the sliding cursor in the second experiment. Yet, numerous real-world stereo systems get this wrong in practice. For example, many games on the Nintendo 3DS mismatch parallax between the cursor and targets. Games using automatic stereo conversion (e.g., NVidia 3D Vision) suffer from the same problem, especially when using the Windows system cursor at zero parallax. Hence, we argue for the need of studies like this, which further analyze the negative impact of what researchers might consider "sloppy" stereo.

We isolate cursor visualization to eliminate these other potentially confounding effects. To isolate this factor, the one-eyed cursor is compared in both stereo *and* mono scene visualizations. The idea of comparing one-eyed and stereo cursors in a mono display scene may initially seem odd. However, our "mono" display actually shows the scene in stereo, but using the same position for both eyes (i.e., 0 disparity). Hence we compare one-eyed and stereo cursors, across screen-plane and sliding techniques in situations where depth should not matter. This allows us to isolate any

negative effects (e.g., discomfort, eye fatigue) of the one-eyed cursor from its benefits (e.g., elimination of depth cue conflicts), allowing us to get at the heart of the possible negative effects of the one-eyed cursor.

2. RELATED WORK

Numerous studies have evaluated 3D trackers as input devices [1]. Studies comparing fish tank VR setups to more immersive systems such as CAVEs [6] or head-mounted displays [11] typically report superior performance with fish tank systems. While fish tank systems often use 3D trackers, they are well suited for mouse-based input due to their desktop-like nature. Moreover, despite numerous available tracker-based 3D selection techniques, mouse input offers better performance if target occlusion is impossible or easily resolved [3, 17].

Head-tracking helps to resolve occlusions and is often used in fish tank VR for head motion parallax cues [2, 6, 11, 20]. Results of experiments on head-tracking tend to be mixed. Some work indicates that head-tracking has a stronger effect than stereo [2], while other work indicates the opposite [20]. We investigate head-tracking to quantify any benefits it offers for mouse-based input. Our rationale is that if head-tracking improves (or at least does not harm) performance, it may be useful in mouse-based systems with target occlusion issues; the user can simply move their head quickly and easily resolve occlusions, without requiring mode changes to use of the mouse for viewpoint control.

Recent work on mouse-based 3D selection [17] showed the benefits of Ware's one-eyed cursor [21] with screen-based cursors. However, one-eyed cursors may cause greater eye strain; recent work instead proposed the use of specialized stereo cursors [12] and found that these offered better performance than the one-eyed cursor. The authors concluded that using more carefully designed stereo cursors, rather than displaying them in the screen plane, can overcome this eyestrain. However, it is difficult to directly compare these results to other 2D pointing techniques. One reason is that the authors do not report common metrics, such as pointing throughput or error rates, and use a non-standard experimental method. Our study aims to add to this body of knowledge by re-evaluating the one-eyed cursor in situations where depth is irrelevant. The objective is to assess any negative impacts of the one-eyed cursor in isolation from situations where it is beneficial, such as when selecting targets at greater depths.

2.1 Fitts' law and Point Selection Tasks

We use a 3D version of the ISO 9241-9 standard [8] based on Fitts' law [7]. The predictive form of Fitts' law is:

$$MT = a + b \cdot ID, \quad \text{where} \quad ID = \log_2\left(\frac{A}{W} + 1\right) \quad (1)$$

MT is movement time and a and b are derived by linear regression. ID is the index of difficulty (in bits). A is the distance to the target (amplitude), and W is the target size. ISO 9241-9 [8] recommends a standardized pointing task (Figure 4) and using *throughput* (TP) as a primary measure:

$$TP = \frac{\log_2\left(\frac{A_e}{W_e} + 1\right)}{MT}, \quad \text{where} \quad W_e = 4.133 \cdot SD_x \quad (2)$$

MT is average movement time, and A_e, (effective amplitude) is the average *actual* movement distance for a given condition. W_e (effective width) is 4.133 standard deviations of the selection coordinates along the target approach vector (SD_x). This adjusts

the error rate to 4%. Throughput has been shown to be consistent despite the inherent speed-accuracy tradeoff in point selection tasks [10]. In contrast, movement time and error rate vary due to the speed-accuracy tradeoff. The main advantage of throughput is its comparative consistency between studies; consequently, it is recommended for inclusion in the analysis of pointing techniques and devices [14]. A 3D extension has been used for 3D pointing experiments [5, 17].

3. ISSUES WITH 2D SELECTION IN VR

This section details the main issues in using a 2D input device (e.g., a mouse) in selecting 3D targets. While these issues most commonly arise in desktop or "fish-tank" VR systems (where a mouse is a convenient input device), the same issues apply in other VR systems as well.

3.1 Stereo Viewing and Diplopia

When selecting a feature displayed at a different depth from the cursor on stereo systems, diplopia (i.e., double vision) occurs. Converging the eyes on the cursor produces two images of the feature. Conversely, converging the eyes on the feature yields two cursor images (Figure 1). The effect becomes more pronounced the greater the depth difference between the cursor and the target. This has been shown to impact 3D selection [17]. A simple solution is to show the cursor only to one eye [21]. While this eliminates the negative impact of diplopia [17], recent work [12] suggests that so-called "one-eyed cursors" may induce greater eye fatigue and may thus perform worse than stereo cursors.
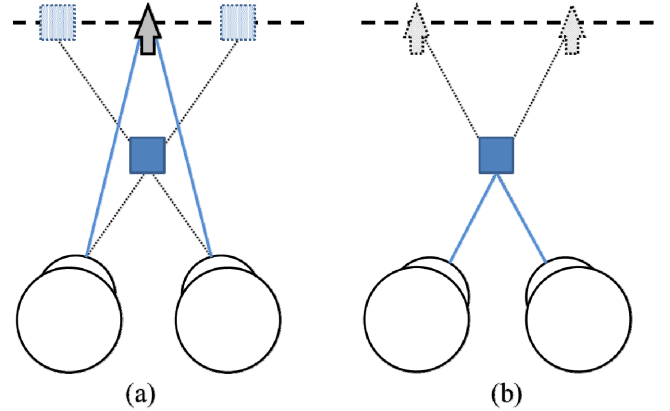

Figure 1. Double-vision. (a) Eyes converging on the cursor. (b) Eyes converging on the feature (a box). The dashed line represents the screen plane.

3.1.1 Sliding vs. Screen-Plane Cursors

Diplopia most commonly occurs with screen-plane cursors [17]. A cursor drawn in the screen-plane (the default position) will generally be presented at a different depth from scene geometry behind it. There are two possible solutions to this problem using different cursor styles. The first simply draws an otherwise normal screen-plane cursor in stereo using the disparity of the surface behind it [12]. The other (which we study) draws the cursor as a 3D object in the scene, and is hence subject to perspective transforms [16, 18]. The cursor is drawn on the foremost visible surface along a ray from the eye position to the screen-plane cursor (which is not shown). Such cursors appear to "slide" along the geometry, hence we refer to them as sliding cursors. An example of sliding cursor motion is depicted in Figure 2.

Initially, one might expect that sliding cursors cannot be occluded, as they are by definition sliding on the visible geometry. Yet, this
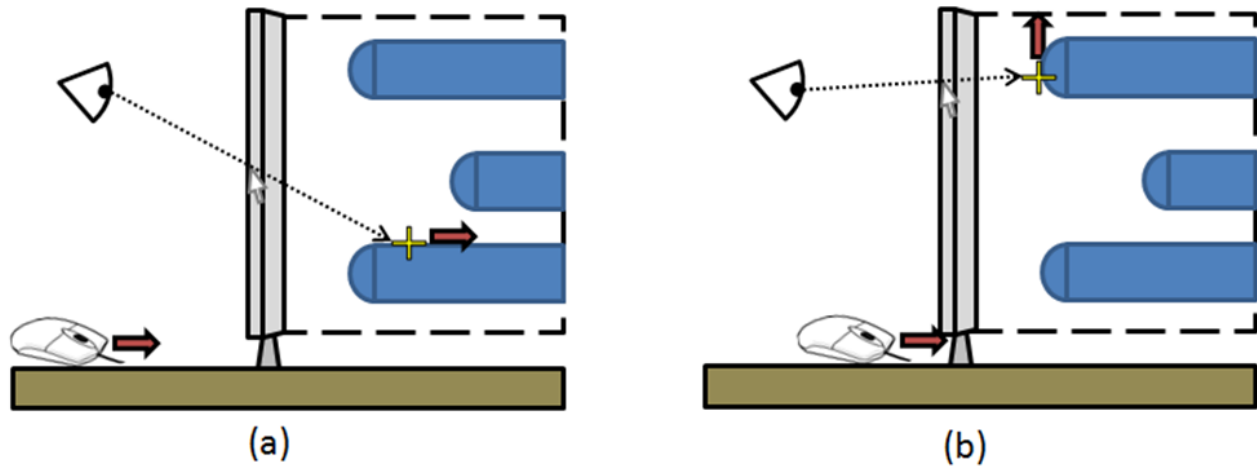
Figure 2. Sliding cursor motion. Dashed box indicates the extents of the virtual scene behind the display (gray box). Blue cylinders represent targets. (a) The position of the cursor (the yellow "+") is determined by the intersection of the eye-mouse ray and the scene. (b) Moving the mouse forward moves the cursor along the manifold of the scene, corresponding to upward motion of the true cursor (which is hidden).

is not completely true, as there are situations when the cursor is visible only from one eye, but not the other. For example, consider a box with one of the side faces aligned with the view direction and straight in front of the viewer. Then the sliding cursor degenerates to a one-eyed cursor. Another issue occurs when the system uses visibility from the virtual "cyclopean" eye in the middle between the two real eyes for determining the sliding cursor position. With this approach, the cursor can even disappear completely, for example, when there is a narrow hole aligned with the view direction straight in front of the viewer. If the cursor is in a hole, it can be invisible from both eyes.

In contrast, screen-plane cursors are the more "traditional" cursors that reside in the screen plane and do not move in depth. Typically, these use the ray from the eye to the cursor for selecting objects in the scene. A straightforward implementation of a stereo screen-plane cursor can be occluded by content displayed (stereoscopically) in front of the screen. An alternative approach is to displays a stereo cursor "on top" of the scene, i.e., in a plane that is closer to the viewer than everything visible in the scene. The camera near plane may be a good candidate, except that this plane may be too close to the viewer for comfortable stereo fusion. Alternatively, the depth buffer can be disabled while drawing the cursor [17]. This ensures that the cursor always occludes the scene geometry, even if the cursor appears in the screen plane. This is the option used in recent work [16] using a one-eyed cursor that was only visible to the dominant eye and moved in the screen plane.

3.2 Transparency vs. One-Eyed Cursor

A final issue investigated in our study is the relative difference due to the visualization of the cursor. Drawing the cursor to only one eye was originally proposed by Ware [21] and has long been used as a simple means of visualizing a 2D cursor in a stereo 3D environment. Recent work [12] suggests that there may be an inherent disadvantage to such a visualization. In particular, a one-eyed cursor may induce greater eye fatigue than a stereo cursor. We investigate this concern further. One issue is that a one-eyed cursor is effectively displayed transparently – since it is only visible to one eye, its opacity is 50% that of a standard stereo cursor. It is unclear if the impact of cursor transparency may be

stronger than the one-eyed visualization. Hence our second experiment includes a transparent visualization (of an otherwise stereo cursor) to investigate the possibility that this transparency rather than eye fatigue is responsible for performance costs.

4. EXPERIMENT 1

The first experiment focused on depth cues issues in mouse-based selection in desktop VR. To this end, we decided to study the influence of stereo display, head-tracking, and cursor visualization (one-eyed vs. stereoscopic) on pointing performance. Stereo display and head-tracking provide additional depth cues, e.g., convergence, stereopsis, and motion parallax. We include the cursor visualization comparison as the one-eyed cursor is necessary for selecting targets away from the screen surface, but may yield (slightly) negative effects [12]. Finally, the experiment also included targets displayed both "in front of" and "behind" the screen surface. This yields both positive and negative parallax situations, and would elicit effects due to stereo display coupled with the cursor visualization.

4.1 Participants

Sixteen participants took part in the study. Their ages ranged from 19 to 39 with a mean age of 23.75 years (SD = 5.13 years). Nine were female. All but two were right-handed, but all used the mouse regularly with their right hand. All participants had normal or corrected-to-normal vision, and were able to see in stereo. This was assessed by showing them our stereo stimulus (a target at 10 cm away from the screen) and asking them to touch where they perceived it to be. Prospective participants who could not find (roughly) the true 3D position of the target were disqualified. All participants were right-eye dominant, determined by a simple thumb occlusion test.

Participant gaming experience was also assessed, as gamers tend to perform better in 3D tasks [13]. Their responses for four game types – mouse & keyboard PC games, console games using a controller, spatial games (using devices such as a Wiimote or Kinect), and mobile games – are summarized in Figure 3. A score of "5" indicates playing every day, while a score of 1 indicates never playing. In general, participants were not regular gamers, indicating at most a score of "several times per month" at most.
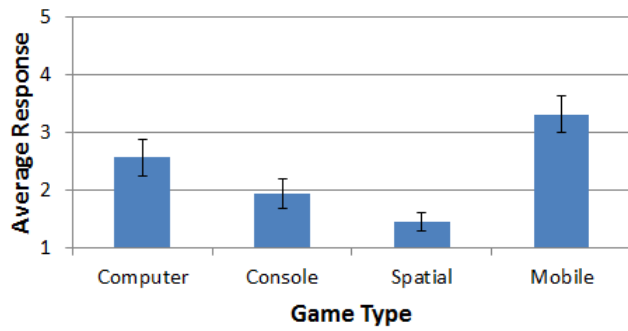
Figure 3. Average gaming experience of participants for Experiment 1. Error bars show ±1 SE.

4.2 Apparatus

The experiment was conducted on an Intel i5-based PC running Windows 7. The PC had a quad-core 3.4 GHz CPU and 16 GB of RAM. An NVidia Quadro 4400 was used with 3DVision Pro glasses for stereo. Software developed in C++ and OpenGL depicted the inside of a wooden crate with target spheres supported on wooden cylinders. See Figure 4. Targets were presented in a circle parallel to the screen surface. All targets in a circle were at the same depth, but depth varied between circles.
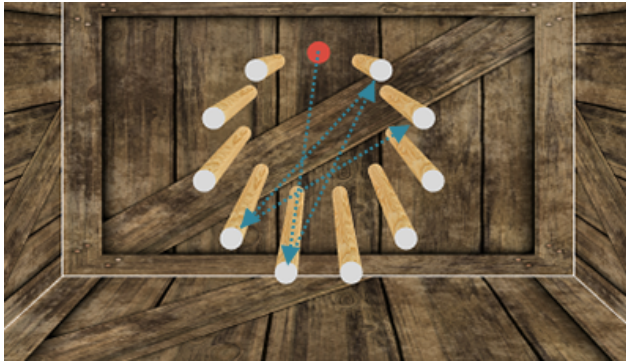

Figure 4. Experimental software depicting 11 targets. Arrows added to illustrate the ordering of the first four targets (not shown in the actual system).

The software could display the scene in stereo or mono, with or without head-tracking, and with a stereo 3D or one-eyed cursor. That cursor was always shown in the screen plane, i.e., at zero parallax. In the mono view condition, the same image was presented (0 disparity) to both eyes and the stereo glasses were active to eliminate the potential confound of different brightness levels. Regardless if the scene was visualized in stereo or mono (noting again that the "mono" display actually used identical left and right eye images), the one-eyed cursor was only shown to the dominant eye, while the stereo cursor was always shown to both eyes. In our system, positive depths correspond to targets in front of the display and negative depths are behind/inside it. When head-tracking was disabled, we used a static viewpoint 65 cm in front of the screen, corresponding to a standard seated position at a desktop workstation. We used a crosshair-shaped cursor that was coloured yellow to stand out against the background. Similar to "real" cursors (e.g., the Windows system cursor), this provides a single and obvious selection hotspot (the centre of the crosshair), which is what most participants are used to.

4.3 Procedure

Participants were first given 20 to 30 practice trials in each condition before starting the actual experiment. They were instructed to select the red target sphere as quickly and accurately as possible, with consistent speed and accuracy. Selection required moving the cursor to the target (projection) and clicking the mouse button.

Following each selection, the target advanced according to the pattern shown in Figure 4 regardless if the target was hit or missed. Upon completing all selections in a target circle, the next circle of targets would appear with different values for target size, distance, and depth. Participants could take breaks when the top target in the circle was active (as in Figure 4), as timing began after that target. The experiment took approximately 1 hour to complete for each participant.

4.4 Design

The experiment used a within-subjects design with the following independent variables and levels:

Stereo Display: Stereo-On, Stereo-Off (i.e., mono)
Head-Tracking: HT-On, HT-Off
Cursor: STC (stereo cursor), OEC (one-eyed cursor)
Target Size: 0.5, 0.75, 1.0 cm
Target Distance: 3.5, 7.5, 9.5 cm
Target Depth: -10, 0, +10 cm

Stereo display, head-tracking, and cursor were counter-balanced with a Latin square. Target size, distance, and depth were selected randomly (without replacement) for each target circle. Target size and distance were not analyzed, and instead only incorporated to create a realistic distribution of task difficulties (per Fitts' law). There were 12 recorded selection trials per target circle. Thus there were $2 \times 2 \times 2 \times 3 \times 3 \times 3 \times 12 = 2592$ trials per participant (i.e., 41472 trials overall). The dependent variables were movement time (ms), error rate (% missed targets), and throughput (bits per second).

4.5 Results

For all dependent variables, head-tracking was not significant and also did not participate in any interaction effects. Since this factor had also very small effect, we collapse its levels (reporting only the averages of HT-on and HT-off) from here on, to help simplify the analysis. We also omit it from the figures below.

4.5.1 Movement Time

Movement time data were not normally distributed ($w = 0.66$, p < .01), which violates one of the prerequisites for ANOVA. Thus, we used Wobbrock's aligned rank transform [22] and performed repeated measures ANOVA on the ranks. There was a significant main effect for target depth on movement time ($F_{2,15} = 4.5$, $p < .05$). The other main effects were not significant. These results must be considered in light of the significant three-way interaction effect between stereo, target depth, and cursor ($F_{2,30} = 33.4$, $p < .00001$). See Figure 5 for movement times.
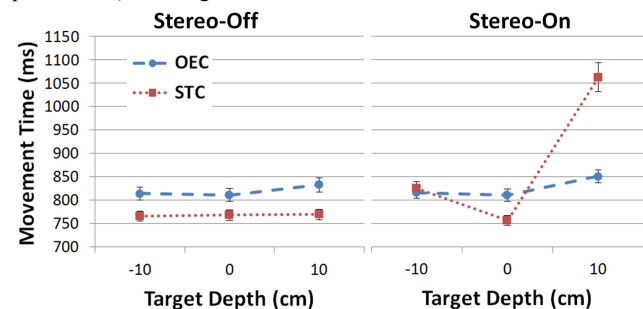

Figure 5. Movement time by condition. Stereo conditions separated by left/right split. Error bars show ±1 SE.

The two conditions with the highest movement time were both STC with stereo-on, at +10 cm target depth, significantly worse than all others. A Tukey-Kramer posthoc test revealed that no other conditions were significantly different ($p < .05$).

4.5.2 Error Rate
Error rate data were not normally distributed ($w = 0.75$, $p < .05$). We analyzed this non-parametrically with ART and repeated measures ANOVA on ranks. Only cursor had a significant main effect ($F_{1,15} = 8.3$, $p < .05$). Yet, the interaction of stereo display, target depth, and cursor was significant ($F_{2,30} = 8.2$, $p < .005$). The combination of stereo-on, STC and +10 cm target depth yielded higher error rates than all other conditions. See Figure 6.
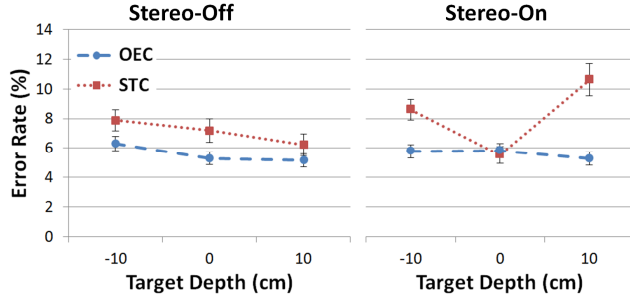

Figure 6. Error rate by condition. Stereo conditions separated by left/right split. Error bars show ±1 SE.

4.5.3 Throughput
We used a variant of "screen-projected" throughput [17]. Rather than projecting targets and selection coordinates to the screen plane, we used the intersection of the mouse ray and target plane as the selection coordinate. A_e and W_e were then computed normally. This effectively projects the task to the target plane and yields the same result in any plane, as throughput depends on the ratio of A_e to W_e. Even though this is a variant, it effectively computes exactly the same value, as the relative over/undershoot of selection coordinates is invariant to the plane where throughput is computed, due to perspective projection (as long as each pair of targets is in the same plane). Note that this variant is inappropriate for scenarios involving varying depth targets. Throughput scores are summarized in Figure 7.
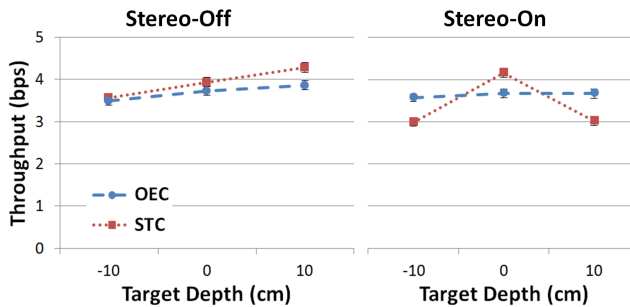

Figure 7. Throughput by condition. Stereo conditions separated by left/right split. Error bars show ±1 SE.

Unlike movement time and error rate, throughput data were normally distributed ($w = 0.95$, $p > .05$) and hence analyzed directly with repeated measures ANOVA. Significant main effects on throughput were found for stereo display ($F_{1,15} = 10.3$, $p < .005$) and target depth ($F_{2,15} = 16.2$, $p < .0001$). There was a significant three-way interaction between stereo, depth, and cursor ($F_{2,30} = 12.4$, $p < .001$).

This interaction is visible in the stereo-on conditions in Figure 7: the two STC 0 cm target depth conditions are significantly higher than the STC +10 cm or -10 cm target depth conditions, but not higher than the OEC conditions. The best stereo-off conditions (both STC at +10 cm target depth) were significantly higher than the stereo-on conditions at both -10 and +10 cm target depths. No other conditions were significantly different. There were no significant differences between any stereo-off conditions.

4.5.4 Subjective Results
We solicited qualitative results from the participants such as preferences for conditions and perceived performance in a questionnaire at the end of the study. Most participants indicated that they felt the one-eyed cursor did not affect their targeting ability (see Figure 8c) and that eye discomfort was not a concern; however, the difference in the number of participants in each response group was not significant ($\chi^2_4 = 7.1$, $p > .05$). Similarly, there was no significant difference in the number of participants in each response group for stereo scene ($\chi^2_4 = 9.0$, $p > .05$) or head-tracking ($\chi^2_4 = 5.9$, $p > .05$).
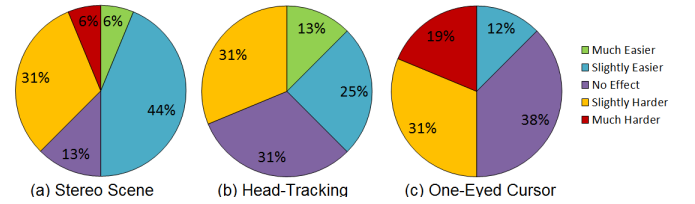

Figure 8. Summary of participant subjective feedback for (a) stereo display, (b) head-tracking, (c) one-eyed cursor. Participants were asked about the perceived effect on targeting for each condition.

Both in the experiment presented here and other work [17] almost all participants without VR/gaming experience were unaware of the presence of the one-eyed cursor – until questioned about it post-experiment. Thus we believe that eye discomfort and similar effects are small concerns, and that a one-eyed cursor can perform as well as a stereo cursor, contrary to previous results [12].

4.6 Discussion
Perhaps the most important finding of this experiment is the relatively small difference between the stereo and one-eyed cursors, contrary to previous work [12]. As suggested by Schemali and Eisemann [12], the one-eyed cursor offers *slightly* worse performance than the stereo cursor condition for conditions where depth does not matter. This includes all stereo-off conditions, and stereo-on conditions at 0 cm target depth. Yet, this effect was not significant. To further investigate this, we conducted a one-way ANOVA using *only* stereo-off conditions across both cursor conditions (stereo and one-eyed). Not only was the effect for cursor *not* significant ($F_{1,15} = 2.7$, $p = .1$), but the effect size was extremely small ($\eta^2 = .01$). The difference in throughput scores was only about 5% (3.9 bps for STC vs. 3.7 bps for OEC). Altogether, this evidence suggests little difference between the mono and stereo cursor visualization in a mono scene.

Thus, we suggest that the most important finding of this experiment is the *absence* of effects for the one-eyed cursor at zero parallax targets. While not definitive proof (since one cannot "prove the null"), this is evidence that the detrimental effects of the one-eyed cursor [12] may well be overstated. Moreover, we specifically designed our experiment to evaluate the effect of the one-eyed cursor in situations where it *should not* make a

difference – if it performed worse, then this would be strong evidence of eye fatigue or discomfort as noted previously [12]. As our experiment failed to identify a significant effect and the differences in throughput are small, we take this as evidence that eye fatigue/discomfort or similar effects are small and that the one-eyed cursor is still a reasonable design choice. Nevertheless, this absence of a statistical effect partly motivates the follow-up experiment presented below. To further explore this issue, the next experiment (again) includes the one-eyed cursor in equivalent conditions to further assess its potential performance impact.

All dependent variables showed significant interactions between stereo display, cursor, and target depth. Like previous work [12, 17], this is the impact of diplopia on targets at different depths from the cursor. When presenting content at drastically different depths from a stereo cursor in a stereo scene, we expect that performance would suffer. The one-eyed cursor largely eliminates this effect, as reflected in the (mostly) flat lines across depth in Figure 5 for the OE cursor. The one-eyed cursor clearly benefits selecting targets at different depths.

Head-tracking did not significantly affect any of the dependent variables. The additional head motion parallax depth cues were not especially helpful with our effectively 2D task, even if it appeared to be 3D. It is possible that this is because participants did not move their heads much after any initial potential exploration of the capabilities of the system and the subsequent discovery of all target positions. Note that head-tracking was not necessary for individual selections. Nevertheless, since performance was also not *worse* with head-tracking, it could be useful to resolve occlusions in environments where occlusions occur frequently enough. Head-tracking also obviates explicit viewpoint mode toggling and helps mouse-based 3D interaction, which might else be hindered by occlusions.

Throughput was consistent with other ISO 9241-9 studies [17]. Previous work [12] that does not report throughput cannot be compared directly to other mouse-based pointing studies due to the speed-accuracy tradeoff inherent in point selection tasks. While the interplay of stereo display and target depth increased throughput variability, these were largely in the expected ~4 bps range. This is especially noticeable in the stereo-off conditions in Figure 7. Contrary to our expectations, performance increased slightly (but not significantly) with closer targets. This may point to a need to re-examine "projected" throughput, even though it eliminated such depth effects in previous work [17].

5. EXPERIMENT 2

We conducted a second experiment to further investigate these issues. This experiment study focused on the interplay between stereo viewing, cursor movement, and cursor visualization.

This experiment was also designed to isolate a previously unexplored difference between the one-eyed and stereo cursor visualizations used in Experiment 1: the effective *transparency* of the one-eyed cursor. Since it is only displayed to one eye, the cursor is effectively 50% transparent relative to the stereo cursor. This experiment thus included a 50% transparent stereo cursor to assess if transparency might have an effect in isolation from displaying the cursor to only one eye.

This experiment also included a sliding cursor [17], implemented as described earlier. The reason we included this condition is that such a cursor is always displayed at the same depth as the target. This avoids issues like stereo cue conflicts and diplopia. Previous

work identified that such cursors perform well [12] with mouse-based selection. This cursor also serves as an approximation of the approach used by modern GPUs when automatically converting 3D content to stereo 3D display. These methods display the cursor in stereo using the disparity of the surface behind it. A main difference between the GPU-based approach and the sliding cursor is that the latter is also subject to perspective and can be occluded by geometry in the scene.

5.1 Participants
Twelve participants took part in the study. Their ages ranged from 20 to 52 years (mean age 26, *SD* of 8.6). Three were female. All used the mouse regularly with their right hand. All participants had normal or corrected-to-normal vision, and were able to see in stereo. This was assessed in the same fashion as in Experiment 1. All participants were right-eye dominant, determined by a simple thumb occlusion test. As in Experiment 1, we assessed their gaming experience. In general, the scores here are higher than Experiment 1; participants of Experiment 2 played games slightly more regularly than those of Experiment 1.
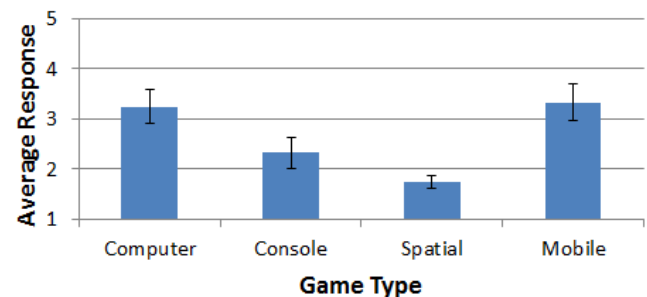

Figure 9. Average gaming experience of participants of Experiment 2. Error bars show ±1 SE.

5.2 Apparatus
Overall, the same apparatus as Experiment 1 was used. Since Experiment 1 indicated that head-tracking had a very limited (if any) effect, we used a static viewpoint in this experiment. We also added two new conditions. The first was a stereo cursor with 50% transparency. The second added a sliding cursor, as described earlier. The sliding cursor was implemented relative to a fixed head position (the "optimal" position, 65 cm from the centre of the screen) as the origin of a mouse ray. The true mouse cursor was not displayed. Instead, a 3D cursor, using the same crosshair shape as in all other conditions, was displayed the intersection point of the mouse ray and the scene. The cursor thus would effectively slide across the scene geometry.

5.3 Procedure and Design
The procedure was identical to that of Experiment 1, with the exception of the two newly added conditions. The experiment used a within-subjects design with the following independent variables and levels:

Stereo Display: Stereo-On, Stereo-Off (i.e., mono)
Cursor: STC, OEC, TRC
Technique: Screen, sliding
Target Size: 0.5, 1.0 cm
Target Distance: 3.5, 7.5, 9.5 cm
Target Depth: -10, 0, +10 cm

The STC and OEC cursors are the same stereo and one-eyed cursor conditions from Experiment 1. The new TRC condition was a stereo cursor displayed with 50% transparency. Technique indicated which cursor control technique was used, either in the
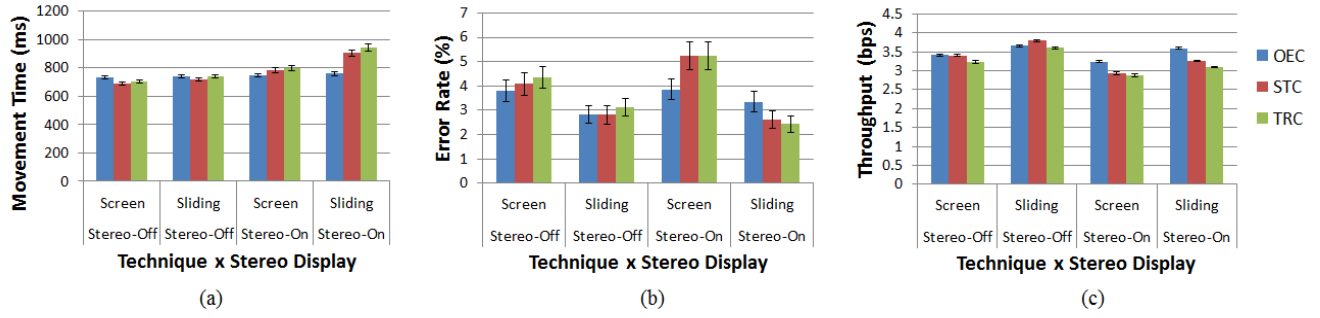
Figure 10. Experiment 2 summary results for (a) Movement Time, (b) Error Rate, and (c) Throughput. These figures show results for each combination of cursor, technique, and stereo display. Target depth is collapsed to simplify presentation. Error bars show ±1 SE.

screen-plane (screen) or geometry-sliding (sliding). Target size and distance were not analyzed, and instead only incorporated to create a realistic distribution of task difficulties (per Fitts' law).

The twelve combinations of stereo display, cursor, and movement were counterbalanced according to a balanced Latin square. The remaining factors (target size, distance, and depth) were presented in random order without replacement.

There were 12 recorded selection trials per target circle. Overall, there were $2 \times 3 \times 2 \times 2 \times 3 \times 3 \times 12 = 2592$ trials per participant, or 31104 trials over all twelve participants. As in experiment 1, the dependent variables were movement time (ms), error rate (% missed targets), and throughput (bits per second).

5.4 Results

With four independent variables of interest, results of this experiment are complex. As a result, we first present an overview/summary of the experiment results for each dependent variable in Figure 10. The figure collapses the target depth factor - i.e., it shows only the cursor, technique, and stereo display combinations, averaging across the three target depths. Figure 11 shows the same information, but in more detail, and separated across the three levels of target depth.

5.4.1 Movement Time

Movement time data were normally distributed ($w = 1.64$, $p > .05$), hence we analyzed results using repeated measures ANOVA. Since the experiment design used four factors (stereo display, cursor, technique, and target depth), and due to the difficulty in analyzing four-way interaction effects, we instead separate the analyses by stereo display. The remaining three factors are analyzed normally (i.e., a 3-way RM-ANOVA).

We first present an analysis excluding depth (i.e., averaging all three depths, and corresponding to the results depicted in Figure 10a). There was a significant interaction effect between stereo display, technique, and cursor ($F_{2,22} = 3.7$, $p < .05$). Tukey-Kramer multiple comparisons revealed that STC and TRC with stereo-on and the sliding cursor were significantly worse than all other conditions (which were not significantly different from each other). As in Experiment 1, results must be interpreted in light of the significant three-way interaction effect between technique, cursor, and target depth ($F_{4,44} = 5.3$, $p < .005$). The effects of diplopia at target depths away from the screen (-10 cm and +10 cm) are visible in Figure 11a for the stereo-on + screen cursor conditions. As before, the one-eyed cursor (OEC) eliminates this effect, and demonstrates fairly flat lines across target depth, as can be seen in Figure 11a. Figure 10a also depicts the consistent performance of OEC across all combinations of conditions. This

is also true for the sliding OEC combination, suggesting little difference between the sliding and screen-based techniques when the one-eyed cursor was used. However, the sliding cursor suffered with the other cursor visualizations–both TRC and STC demonstrate significantly worse performance the higher the target depth (i.e., the *closer* the target is to the viewer). In fact, the worst condition overall was the combination of a transparent sliding cursor, at +10 cm target depth, with stereo-on.

We also investigate separate main effects to help better explain the interaction effect noted above. For stereo-off conditions, there was a significant main effect for cursor ($F_{2,11} = 4.5$, $p < .05$). Tukey-Kramer posthoc analysis revealed that the OEC conditions were significantly slower than either the TRC and STC conditions ($p < .05$). This effect was more strongly pronounced for the sliding cursor conditions. See Figure 10a and Figure 11a. For stereo-on conditions, the results are more complex. There were significant main effects for technique ($F_{1,11} = 32.5$, $p < .0005$), cursor ($F_{2,11} = 6.9$, $p < .005$), and target depth ($F_{2,11} = 41.5$, $p < .0001$). This is especially noticeable in Figure 10a: the sliding cursor with stereo-on offered much worse performance with both STC and TRC than with OEC. Figure 11a highlights that this is due to the impact of different target depths, suggesting the effect is again, due to diplopia.

5.4.2 Error Rate

Error rates were normally distributed ($w = 3.6$, $p > .05$) hence we analyzed results using repeated measures ANOVA. As with movement time, we separated the error rate analysis by the stereo display factor. Error rates are summarized in Figure 10b, and presented in greater detail in Figure 11b.

We first analyzed error rates by averaging across target depth. Only the main effect for technique was significant ($F_{1,11} = 33.5$, $p < .0001$), with the sliding cursor globally offering lower error rates than the screen cursor. This is likely due to the high variability in error rates, particularly noticeable in Figure 11.

To analyze the effects of depth, we consider the stereo-on and stereo-off conditions separately (to avoid high-order interaction effects). For stereo-off conditions, the main effect for target depth was significant ($F_{2,11} = 3.9$, $p < .05$). Farther target depths (-10 cm) offered significantly higher error rates at about 4.1% than either of the other two target depths at 3.1% for +10 cm targets and 3.3% for 0 cm targets. The main effect for technique was also significant ($F_{2,11} = 13.5$, $p < .005$), with the sliding cursor offering better (on average) error rates than the screen cursor. No other main or interaction effects were significant for stereo-off conditions.
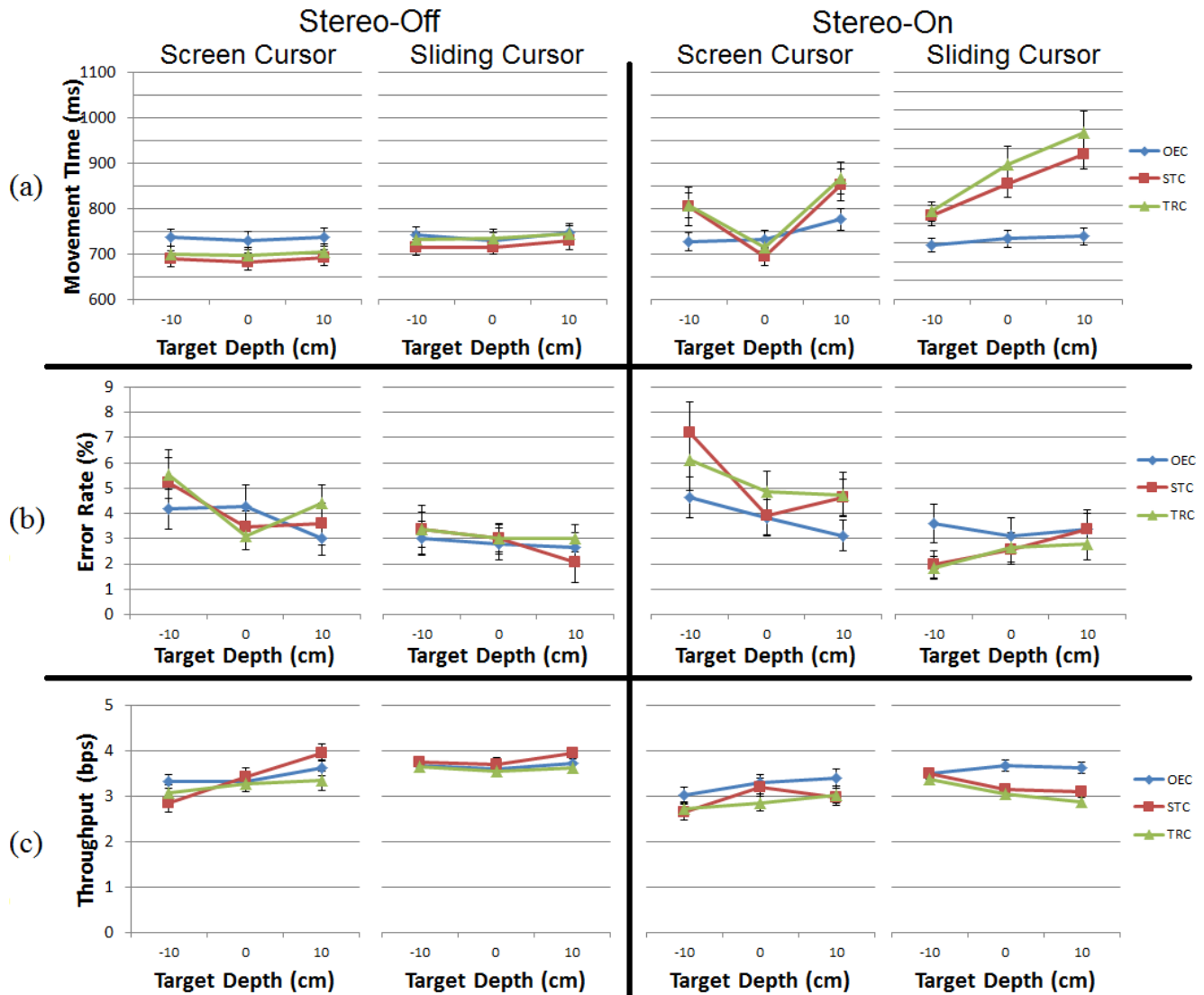
Figure 11. Experiment 2 results for (a) Movement Time, (b) Error Rate, and (c) Throughput. The conditions on the left side are all stereo-off (mono) conditions, while the right side shows stereo-on conditions. Within these halves, the left side shows the screen-based cursor technique and the right side shows conditions using the sliding cursor technique. Error bars show ±1 SE.

For stereo-on conditions, the only significant effect was the main effect for technique ($F_{1,11} = 21.7$, $p < .001$). Neither cursor ($F_{2,11} = 0.21$, ns), nor target depth ($F_{2,11} = 0.71$, ns) were significantly different, nor were any interaction effects. Overall, the sliding technique offered substantially better error rates than the screen-based technique, 2.8% vs. 4.8% respectively. This is likely a side effect of the slower movement time with this technique: Participants took their time to ensure careful selection, likely due to the comparative difficulty in using the technique. Note that relative absence of statistical effects for error rate is likely due to the comparatively high degree of variability in error rates (unlike movement time). For example, diplopia appeared to negatively impact error rates with the screen-based technique, particularly with the STC condition at -10 cm. However, because of the high variability, this effect was not found to be significant.

5.4.3 Throughput
Throughput was calculated as described earlier – using the relative over/undershoots in the target plane, rather than the screen

plane. This yields the same score as a screen-space calculation [17]. Throughput scores were normally distributed ($w = 1.6$, $p > .05$) hence we analyzed results using repeated measures ANOVA. Throughput scores are summarized in Figure 10c and Figure 11c.

As with the other dependent variables, we first analyze only the "simplified" result, averaging throughput scores over target depth for each condition. See Figure 10c. Significant main effects included stereo display ($F_{1,11} = 59.7$, $p < .001$), technique ($F_{1,11} = 12.6$, $p < .005$), and cursor ($F_{2,11} = 9.6$, $p < .001$). However, the technique scene/cursor interaction effect was also significant ($F_{2,22} = 4.8$, $p < .05$). Notably, the combination of stereo-on, screen cursor, and TRC was the worst performer, significantly worse (via Tukey-Kramer, $p < .05$) than all stereo-off conditions, as well as the corresponding OEC condition. The remaining conditions were not significantly different.

As with the other dependent variables, we independently analyzed stereo-on and stereo-off conditions to simplify analysis. We first report the results for stereo-off conditions. There were significant

main effects for technique ($F_{1,11} = 8.3$, $p < .01$), and target depth ($F_{2,11} = 8.1$, $p < .005$). The interaction effects between technique and target depth was significant, ($F_{2,22} = 6.3$, $p < .01$), as was the interaction effect between cursor and target depth ($F_{4,44} = 2.9$, $p < .05$). These effects are largely visible in Figure 11c (depicted as crossing lines). Notably, the OEC and TRC performance was consistent across depth, while STC offered significantly higher performance with *closer* targets using the screen-based cursor. Performance was much more consistent with all cursors for the sliding cursor.

For stereo-on conditions, there were significant main effects for technique ($F_{1,11} = 12.7$, $p < .005$), and cursor ($F_{2,11} = 12.8$, $p < .0005$). On average, the one-eyed cursor (OEC) offered significantly higher throughput at 3.4 bps vs. 2.98 bps for TRC and 3.1 bps for STC. The slide technique also offered higher throughput than the screen cursor. However, there was also a significant interaction effect between technique and target depth ($F_{2,22} = 5.1$, $p < .05$). This is visible in Figure 11c. The sliding cursor performed worse with closer (+10 cm) targets, except with the OEC cursor, which was largely immune to this effect. This may be because participants tended to slide the TRC and STC cursors up the sides of target cylinders. With OEC, they were most likely unaware they were using the sliding cursor, and hence treated it like a "normal" mouse cursor (i.e., not sliding).

5.4.4 Subjective Results

We also solicited subjective feedback from participants about their experience with the various conditions. In particular, we asked participants if they felt the stereo display and the one eyed cursor made targeting easier or harder. A significant number of participants felt stereo display made targeting slightly harder ($\chi^2_4 = 18.8$, $p < .001$). There was no significant difference in the number of participants in each response group for the one-eyed cursor though ($\chi^2_4 = 4.7$, $p > .05$). See Figure 12.
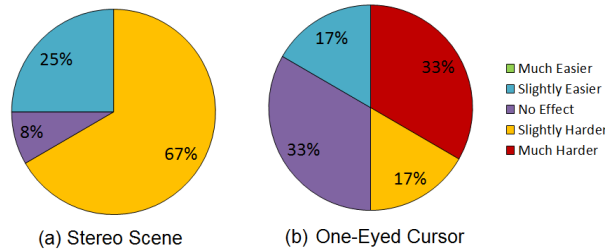
Figure 12. Summary of participant subjective feedback for (a) stereo vs. mono display, (b) one-eyed cursor vs. stereo cursor. Participants were asked about the perceived effect on targeting for each condition.

We were also interested in their perceived performance with both the sliding cursor technique (compared to the screen-based technique) and whether they were aware of the difference between the one-eyed and transparent cursors. They were asked "*I found that the sliding cursor improved my targeting ability relative to the screen cursor.*" and "*I was able to tell the difference between the one-eyed and transparent cursors*". These responses were solicited on a 5-point Likert scale ranging from "Strongly Disagree" to "Strongly Agree". Figure 13 summarizes the percentage of participants in each response group. The number of participants giving each response was not significantly different for the sliding cursor ($\chi^2_4 = 3.8$, $p > .05$). Similarly, participants were unable to reliably tell the difference between the one-eyed and transparent cursors ($\chi^2_4 = 0.5$, ns) - there was a roughly equal number of participants giving each response.
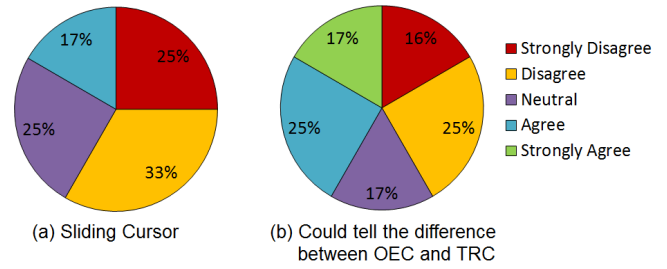
Figure 13. Subjective responses for (a) if they felt the sliding technique offered better performance than the screen technique, and (b) if they could tell the difference between TRC and OEC.

5.5 Discussion

For the stereo-off conditions (i.e., mono display), the difference between the various cursor conditions was minimal. This matches the outcome of a previous comparison of the one-eyed and stereo cursors [17]. However, the combination of the one-eyed cursor and screen-plane technique showed slightly (and significantly) slower performance than the others. Note that in this condition the one-eyed cursor should have no effect *at all*; this may support previous work indicating negative aspects of the one-eyed cursor [12]. However, the difference is only significant in terms of movement time, but not in throughput. The transparent cursor also did not behave differently than the stereo cursor.

For the stereo-on conditions with the stereo cursor (STC), the sliding cursor performed worst. More interestingly, the transparent cursor performed significantly worse than the one-eyed cursor (but similar to the stereo cursor). This makes sense, as it was effectively a slightly *worse* stereo cursor. This is most noticeable at or in front of the screen – closer targets reduced performance with the sliding technique. Performance with the screen-plane cursor was worse away from the screen for the STC and TRC conditions. However, the screen plane cursor also offered worse error rates for deeper targets, potentially due to the impact of perspective - targets appeared smaller, and thus harder to select reliably [17]. This is clearly due to the effect of diplopia - as expected, this is most pronounced for targets farther from the screen, but only applies in stereo scenes. Surprisingly, the sliding cursor did much worse. This may be because participants tended to slide the cursor up the fronts of cylinders, as noted in previous work [17], despite the ability to simply move the cursor over the target (i.e., it would "pop" to the front).

It is also worth noting that the one-eyed cursor "leveled" performance across all conditions. As seen in Figure 11, the movement time and throughput lines for the OEC conditions are essentially flat across all conditions. This is an interesting result, and suggests limited differences between these conditions when the cursor is visualized in mono. Even the difference between the sliding and screen techniques was small.

Overall, and contradicting previous work, the one-eyed cursor performed well in situations where it was expected to – i.e., to eliminate negative impact of stereo cue conflicts in the stereo-on conditions. However, our study provides additional evidence that this visualization may yield some fatigue or similar effects, as previously suggested [12]. The cursor performed comparatively (slightly) *worse* in mono scene visualization (stereo-off) than stereo cursors – a condition where the benefits of the cursor are isolated from any potential negative impact it may have. However, we note that this is an unusual and unrealistic combination that would be rarely, if ever, used in practice. Thus,

we can still overall recommend the one-eyed cursor as a reasonable design choice in desktop 3D systems. Ultimately, the OEC moving in either the screen-plane or sliding on geometry offered better performance than all other options in terms of movement time and throughput for common user interface setups, such as stereo display.

6. CONCLUSION

We presented two experiments on issues in on mouse-based 3D selection in desktop VR systems. The first experiment investigated stereo, head-tracking, and target depth. The study included the one-eyed cursor, to mitigate the negative effects of stereo conflicts commonly observed with mouse cursors in stereo 3D displays. The second experiment investigated if any negative effects of the one-eyed cursor may be due to its comparative transparency, while also comparing a screen-based and geometry-sliding cursor. Overall results confirm that presenting the cursor in stereo significantly impacts performance for different depth targets due to diplopia. The one-eyed cursor eliminates this effect, performing significantly better than any stereo cursor techniques when the scene is visualized in stereo. However, confirming previous findings, there is a small negative impact of the one-eyed cursor visualization in *mono* scenes, i.e., where its positive effects are ruled out. Consequently, it is reasonable that alternative stereo cursor presentation methods may yield better performance still. Further study on novel cursor techniques for desktop VR is still warranted.

7. ACKNOWLEDGEMENTS

8. REFERENCES

[1] Argelaguet, F. and Andujar, C., A survey of 3D object selection techniques for virtual environments, *Computers & Graphics*, *37*, 2013, 121-136.

[2] Arthur, K. W., Booth, K. S., and Ware, C., Evaluating 3D task performance for fish tank virtual worlds, *ACM Transactions on Information Systems*, *11*, 1993, 239-265.

[3] Bérard, F., Ip, J., Benovoy, M., El-Shimy, D., Blum, J., and Cooperstock, J., Did "Minority Report" get it wrong? Superiority of the mouse over 3D input devices in a 3D placement task, in *Human-Computer Interaction – INTERACT* 2009, 400-414.

[4] Bogdan, N., Grossman, T., and Fitzmaurice, G., HybridSpace: Integrating 3D freehand input and stereo viewing into traditional desktop applications, *Proc. of the IEEE Symposium on 3D User Interfaces - 3DUI 2014*, (New York: IEEE, 2014), 51-58.

[5] Bruder, G., Steinicke, F., and Sturzlinger, W., To touch or not to touch? Comparing 2D touch and 3D mid-air interaction on stereoscopic tabletop surfaces, *Proc. of ACM Symposium on Spatial User Interaction - SUI 2013*, (New York: ACM, 2013), 9-16.

[6] Demiralp, C., Jackson, C. D., Karelitz, D. B., Zhang, S., and Laidlaw, D. H., Cave and fishtank virtual-reality displays: A qualitative and quantitative comparison, *IEEE Transactions on Visualization and Computer Graphics*, *12*, 2006, 323-330.

[7] 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.

[8] ISO, ISO 9241-9 Ergonomic requirements for office work with visual display terminals (VDTs) - Part 9: Requirements for non-keyboard input devices: International Standard, International Organization for Standardization, 2000.

[9] Kulshreshth, A. and LaViola Jr, J. J., Evaluating performance benefits of head tracking in modern video games, *Proc. of the ACM Symposium on Spatial User Interaction - SUI 2013*, (New York: ACM, 2013), 53-60.

[10] MacKenzie, I. S. and Isokoski, P., Fitts' throughput and the speed-accuracy tradeoff, *Proc. of the ACM Conference on Human Factors in Computing Systems - CHI 2008*, (New York: ACM, 2008), 1633-1636.

[11] Qi, W., Taylor II, R. M., Healey, C. G., and Martens, J.-B., A comparison of immersive HMD, fish tank VR and fish tank with haptics displays for volume visualization, *Proc. of the Symposium on Applied Perception in Graphics and Visualization*, (New York: ACM, 2006), 51-58.

[12] Schemali, L. and Eisemann, E., Design and evaluation of mouse cursors in a stereoscopic desktop environment, *Proc. of the IEEE Symposium on 3D User Interfaces - 3DUI 2014*, (New York: IEEE, 2014), 67-70.

[13] Smith, S. P. and Du'Mont, S., Measuring the effect of gaming experience on virtual environment navigation tasks, *Proc. of the IEEE Symposium on 3D User Interfaces - 3DUI 2009.* (New York: IEEE, 2009), 3-10.

[14] 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.

[15] Teather, R. J. and Stuerzlinger, W., Exaggerated head motions for game viewpoint control, *Proceedings of the ACM Conference on Future Play 2008*, (New York: ACM, 2008), 240-243.

[16] Teather, R. J. and Stuerzlinger, W., Pointing at 3D targets in a stereo head-tracked virtual environment, *Proc. of the IEEE Symposium on 3D User Interfaces - 3DUI 2011*, (New York: IEEE, 2011), 87-94.

[17] Teather, R. J. and Stuerzlinger, W., Pointing at 3D target projections using one-eyed and stereo cursors, *Proc. of the ACM Conference on Human Factors in Computing Systems - CHI 2013*, (New York: ACM, 2013), 159 - 168.

[18] Teather, R. J. and Stuerzlinger, W., Visual aids in 3D point selection experiments, *Proceedings of the ACM symposium on Spatial user interaction - SUI 2014*, (New York: ACM, 2014), 127-136.

[19] Wang, G., McGuffin, M. J., Berard, F., and Cooperstock, J. R., Pop-up depth views for improving 3D target acquisition, *Proc. of Graphics Interface 2011*, (Toronto: CIPS, 2011), 41-48.

[20] Ware, C. and Franck, G., Evaluating stereo and motion cues for visualizing information nets in three dimensions, *ACM Transactions on Graphics*, *15*, 1996, 121-140.

[21] 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.

[22] Wobbrock, J. O., Findlater, L., Gergle, D., and Higgins, J. J., The aligned rank transform for nonparametric factorial analyses using only anova procedures, *Proc. of the ACM Conference on Human Factors in Computing Systems - CHI 2011*, (New York: ACM, 2011), 143-146.