Visual Aids in 3D Point Selection Experiments

Robert J. Teather
McMaster University
1280 Main Street West
Hamilton, Ontario, Canada
teather@mcmaster.ca

Wolfgang Stuerzlinger
York University
4700 Keele Street
Toronto, Ontario, Canada
wolfgang@cse.yorku.ca

ABSTRACT

We present a study investigating the influence of visual aids on 3D point selection tasks. In a Fitts' law pointing experiment, we compared the effects of texturing, highlighting targets upon being touched, and the presence of support cylinders intended to eliminate floating targets. Results of the study indicate that texturing and support cylinders did not significantly influence performance. Enabling target highlighting increased movement speed, while decreasing error rate. Pointing throughput was unaffected by this speed-accuracy tradeoff. Highlighting also eliminated significant differences between selection coordinate depth deviation and the deviation in the two orthogonal axes.

Author Keywords

3D pointing; selection; cursors; stereo display; depth cues.

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.

General Terms

Performance, Human Factors.

INTRODUCTION

Target selection is a fundamental task in 3D user interfaces, as in the 2D desktop paradigm. However, 3D selection is considerably more complicated and comparatively less well-understood, despite considerable research on the topic [3, 4, 14]. Perhaps the largest difference is that selecting a point in a 3D volume requires control of at least three degrees of freedom (3DOF). Virtual hand techniques, for example, require movement in each of the x, y, and z axes. In contrast, 2D selection requires only 2DOFs, which is readily handled by the mouse. To date, no standard 3D input device or selection technique exists.

Systems that employ 3D selection typically also use 3D display technologies, which introduce additional complexities. Targets farther from the viewer appear visually smaller due to the perspective transformation. Yet, it is well-known that target size affects selection difficulty [15, 23]. The influence of perspective-scaled target size largely depends on the selection technique [37]. Current stereo displays also introduce the well-known conflict between vergence and accommodation [17]. Consequently, selection of targets in 3D space (e.g., via direct touch) is difficult [11, 36], even with the additional depth cues afforded by stereo.

Researchers investigating 3D selection often add additional visual aids to improve performance. Extra depth cues such as head motion parallax [20], texturing [34], or shadows [18] enhance perception of depth, and may thus improve selection performance. Other forms of feedback may also be helpful. Haptic feedback, for example, helps participants "feel" target depths, which may improve performance [12]. Its absence, however, may impair one's ability to find the true depth of targets [36].

Recent work [11, 36] employed specific visual aids for 3D touch-based selection experiments. Aside from stereo and head-tracking, these aids include texture, avoiding floating targets, and "highlighting" for extra visual feedback – i.e., changing target colour when touched/intersected by the tracker/finger. These studies used a 3D extension of the ISO 9241-9 standardized methodology, which improves comparability between studies [19]. However, it is unclear if such visual aids influence the consistency of results. For example, is it valid to compare results of a 3D selection study using texturing to one that does not?

We address this methodological concern by evaluating the impact of each factor on 3D point selection. We present a study quantifying the effects of texturing, target highlighting, and "support cylinders" that eliminate floating objects. The rationale behind avoiding floating objects is that such objects are rare in the real world. Humans may have a hard time with such an atypical task. Our experiment uses the same 3D variant of the ISO 9241-9 method employed by others [11, 36] . We exclusively used touch-based interaction: a tracked stylus that required participants to touch the tip to targets in 3D to select them.

RELATED WORK

Most direct 3D selection techniques fall roughly into two broad paradigms: ray-based techniques (including occlusion) and virtual hand techniques [1, 7, 14, 28]. Virtual hands use intersection of the hand/cursor with the

SUI'14, October 4–5, 2014, Honolulu, HI, USA.

ACM 978-1-4503-2820-3/14/10…$15.00.
http://dx.doi.org/10.1145/2659766.2659770

target and thus require depth precision. Since our current work focuses on evaluating visual aids in depth motions, we exclusively study virtual hands. Our virtual "hand" uses a tracked stylus to approximate the actual hand position.

Virtual hand (and ray-based) selection techniques are largely equivalent to pointing tasks, i.e., they specify a unique position (of an object) in the environment. Numerous 3D pointing studies have been conducted [4, 12, 21, 27, 35], yet 3D pointing is still not as well understood as its 2D equivalent [31]. Research comparing pointing in the real world vs. virtual reality indicates that VR performance is substantially worse [21, 26, 32]. Several factors contribute to this difference, notably including input latency and noise [35], tracker registration [32] and tactile feedback [12]. However, visual cues seem critical as the largest differences occur during the correction phase of motion, where visual feedback is used in a tight feedback loop [21, 26].

Due to the direct correspondence between input and display spaces, target selection is affected by several visual cues and feedback mechanisms. Early work [4, 5] focused on stereo and head-tracking and found that target position significantly affected task completion time and accuracy. Depth movements were slower and less accurate than screen-parallel movements. Participants were better able to judge depth with stereo enabled. These results were later confirmed in a docking task [5] – stereo significantly reduced movement error in depth.

Other visual aids are important in 3D pointing. Partial target occlusion improves selection with volumetric cursors, especially when combined with stereo [39]. Visual feedback also improves object position memorization [13]. Color change is a commonly used visual feedback mechanisms. The recent Virtual Mitten technique [1], for example, uses colour changes to indicate pressure applied to a handheld grip device. Other recent work focused on visual feedback for hand-based grasp techniques [29]. The authors report that changing the selected object colour was preferred by participants, even though it did not necessarily offer the best performance. Similar approaches improved participant speed and accuracy of 2D pointing in sub-optimal viewing conditions [16]. However, highlighting selected targets in a 2D pursuit tracking task did not improve performance [25].

Pointing and Fitts' Law

We use a 3D extension of the ISO 9241-9 standard [19] based on Fitts' law [15]. This paradigm has been previously employed in 3D pointing studies [10, 11, 36, 37].

Fitts' law states that the difficulty in selecting an object is based on the distance to and size of the target. Increasing distance or decreasing size increases selection time and vice versa. The predictive form of the model is thus given by:

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

MT is movement time, and a and b are empirically determined via linear regression for a given condition. ID is the index of difficulty (in bits), which indicates overall task difficulty. A is distance to the target (amplitude), and W is the target width.

ISO 9241-9 [19] employs a standardized pointing task (Figure 1) based on Fitts' law. The standard uses throughput as a primary measure [6]. Throughput (TP) is defined in bits per second as:

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

The log term is the *effective* index of difficulty, ID_e, and MT is the average movement time for a given condition. Effective amplitude, A_e, is the average *actual* movement distance for a given condition. Effective width (W_e) is computed using SD_x, the standard deviation of the distances between the selection coordinates and the target along the task axis (the line between adjacent targets). By multiplying SD_x by 4.133 (±2.066) standard deviations, the experiment accuracy rate is adjusted to 96%, i.e., to a 4% miss rate. This accuracy adjustment enables the comparison between studies with differing error rates [23].

By "normalizing" experimental accuracy, throughput incorporates speed and accuracy into a single measure. It has been shown to be unaffected by speed-accuracy trade-offs [24]. Effective measures better account for user behavior, and thus enhance comparison.

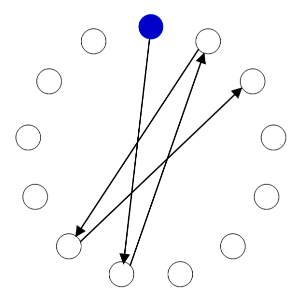

Figure 1. ISO 9241-9 reciprocal selection task with thirteen targets. Participants click the blue target, starting with the top-most one. Targets advance in the pattern indicated by the arrows. Arrows show the ordering for the first four targets.

ISSUES IN 3D POINTING STUDIES

Several issues arise in extending the ISO standard to 3D pointing studies. It is unclear how important each is in ensuring consistent results. While some issues have been identified before [33], we review the most important ones here. Note that additional issues may arise in general 3D point selection studies.

Stereo Viewing

A major concern is cue conflicts inherent to stereo displays. The human visual system unable to focus simultaneously at objects at different depths (e.g., a finger and a target), and most stereo systems suffer from the vergence-accomodation conflict. Consequently, when focusing on a 3D target displayed on the screen, viewers see a blurred finger or when focusing on the finger, they see a blurred target [9]. Systems using stereo touchscreens also suffer from diplopia [38]. This likely impacts both the initial ballistic phase of pointing, as the motor program may target the wrong location in space, and also the correction phase, where visual cues are very important [21, 26].

Cursor/ray-based selection avoids full 3D pointing, as they only select objects visible from the viewer or along a ray. Thus, stereo may be less important with such techniques than virtual hands, but likely still improves scene perception. A 2D performance model describes such techniques quite well [37]. Displaying the cursor in mono eliminates (at least some of) the negative effects of stereo conflicts, yet may cause greater eye strain [30]. We exclusively investigate stylus-based touch in our experiment; pilot testing revealed that stereo is necessary for such techniques.

Selection Distribution
Another concern is that the distribution of selection coordinates in a 3D pointing experiment may not be spherical [22, 40]. This may be due to depth perception inaccuracies. This is illustrated by a recent analysis of 3D touch on a tabletop [11, Fig. 6]. The definition of throughput in ISO9241-9 relies on a (roughly) symmetrical and normal hit distribution [19]. Large deviations may invalidate the underlying assumption(s) of the accuracy adjustment. Thus, part of our work focuses on investigating the distribution of selection coordinates in a 3D pointing task.

Floating Targets
Objects floating in space correspond poorly to real world pointing tasks. Fitts' original pointing experiments, for example, involved selecting physical objects in the real world with a pen [15]. Consequently, to enhance realism in pointing experiments, some researchers [21, 36] have used objects (e.g., cylinders) as "support pedestals" for targets. This helps visually "anchor" the targets in space. However, it is unclear if this is necessary; the effect of such cylinders has not been previously evaluated.

Target Shapes
Another question concerns the shape of the target area or volume participants must select. The two most obvious choices are discs and spheres. Discs are equivalent to 2D targets. This enables direct comparisons between 2D and 3D pointing [36]. However, the visual profile of discs depends on the viewing angle. Consequently, they make the most sense when oriented "flat" toward the viewer. Their extremely small depth extent also makes them difficult to select with touch-based techniques; it is unlikely that participants are able to reliably intersect a flat disc target. For discs, a crossing paradigm [2] is likely more appropriate, but would prevent comparison with other pointing studies.

Spherical targets, on the other hand, are the natural 3D extension of the circular targets recommended by ISO9241-9. Like circles, they have a single "size" parameter, i.e., diameter. A disadvantage is that positioning spheres on top of "support pedestals" may be visually strange. An alternative option is to use a hemisphere, which effectively centers the sphere at the top of the cylinder. This option may distort the computation of effective target width, though. In our experiment, we opt to use (full) spherical targets.

Selection Feedback
Several cues indicate when we have touched a target in reality, including tactile feedback and stereo viewing with correct vergence and accommodation. However, most VR systems do not present these correctly, if at all. Consider, for example, selecting a 3D target using a tracked finger. Due to the absence of tactile feedback, the finger will pass *through* the target. Stereo cues now indicate that the target is in front of the finger, while occlusion cues indicate the opposite – the finger *always* occludes the screen. Consequently, another means of selection feedback is required.

Recent work [11, 36] used target highlighting for this very reason. When the target is touched, it changes colour. This provides feedback that selection (e.g., via a button) will be successful, and helps the user choose between multiple targets. Our current work addresses the question of how important this feedback really is.

METHODOLOGY
Participants
Sixteen participants took part in the study (aged 19 to 39, mean age of 23.4, SD = 4.5 years). Eight were female. All had normal stereo vision. Stereo vision was assessed by showing a stereo stimulus 10 cm in front of the screen and asking them to touch its perceived position. Participants were disqualified if they could not (roughly) find the object's 3D position.

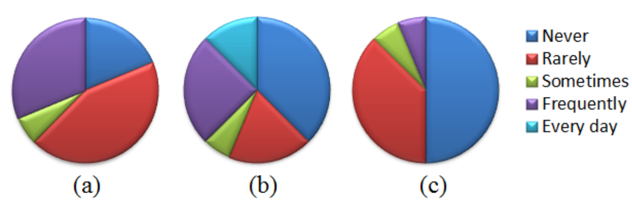

Figure 2. Participant gaming with (a) mouse and keyboard games, game controllers, and spatial devices (Wiimotes, Kinects, etc.). Rarely: once or twice per month. Sometimes: several times per month. Frequently: several times per week.

Participants were asked about their gaming experience. Results are summarized in Figure 2 for mouse/keyboard games (Figure 2a), console controller games (Figure 2b), and spatial input games, e.g., using Wiimotes, Kinects, or similar devices (Figure 2c).

Apparatus

Hardware

The experiment was conducted using a PC running Windows 7. The PC had an Intel *i5* quad-core 3.4 GHz processor, 16 GB of RAM. An NVidia Quadro 4400 graphics card that supports quad-buffered stereo graphics was used, with NVidia 3DVision Pro glasses for stereo. The setup is depicted in Figure 3.

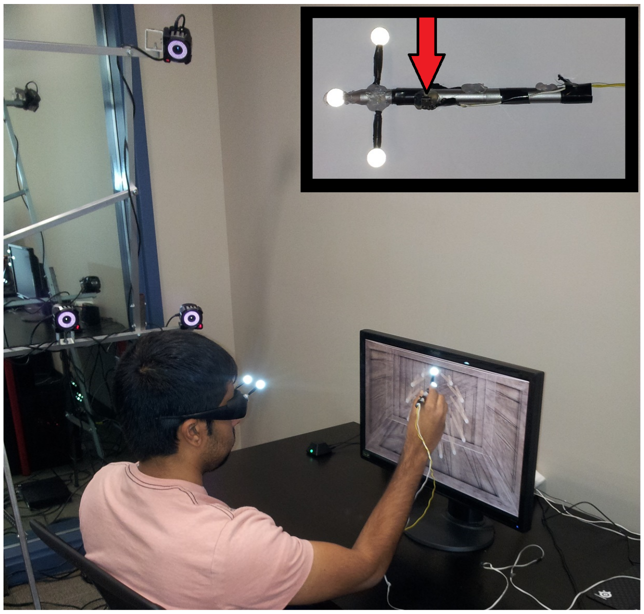

Figure 3. A participant performing the task in our fish-tank VR setup. Top right inset: Stylus used in the experiment with tracker markers. The stylus button is indicated by an arrow.

A tracked stylus was used as the input device. The marker positioned at the tip of the stylus was treated as the "hot spot". The stylus used a re-engineered USB mouse to provide button click events. The button was positioned where the thumb would naturally grip the stylus, i.e., opposite to the fingers. See Figure 3 inset. This was intended to reduce or eliminate the so-called "Heisenberg effect" [8] – that pressing a button on a tracked input device can cause the tip to move slightly upon selection.

The glasses and stylus were tracked using five NaturalPoint *Optitrack S250e* cameras calibrated to ~0.5 mm precision with NaturalPoint's software. We did not use smoothing, as the increased latency may outweigh the benefits [35].

Software

Software was developed in C++ and OpenGL and depicted the inside of a box. See Figure 4. The scene was always presented in stereo with head-tracking as the task was impossible without these cues. Targets were presented as spheres. Depending on the condition, targets either floated at or in front of the display, or were positioned at the tops of support cylinders, see Figure 4. Targets were always presented in a circle parallel to the display surface, like (most) previous work. The software logged selection times, error rates, selection coordinates, and stylus motion trails.

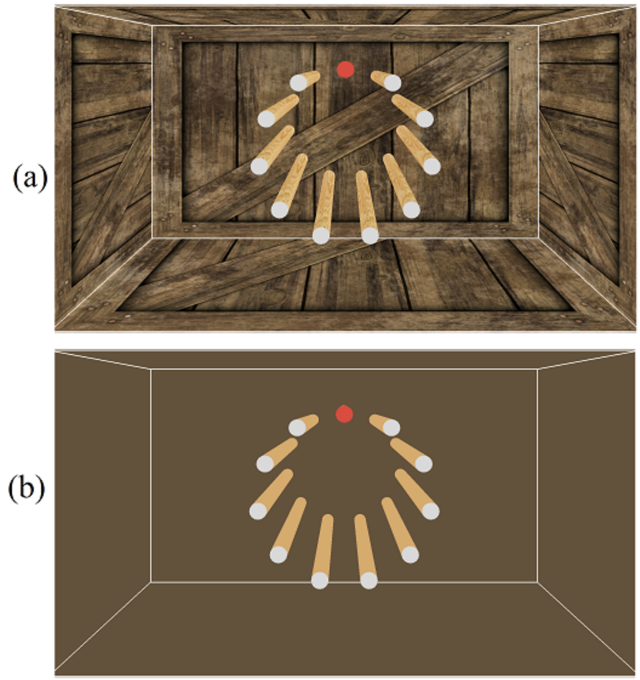

Figure 4. The scene presented by the software (a) Texture-on view. (b) Texture-off view. The "wooden" cylinders supporting the target spheres were excluded in the "cylinders-off" conditions. This figure shows targets at the screen surface.

Similarly, depending on the condition the entire scene was either presented with or without textures. A wood-grain texture with strong texture gradient was used on the cylinders, while a wooden crate texture was used on the box (Figure 4a). In the texture-off conditions (Figure 4b), the cylinder and box were set to the average colour of their respective texture to maintain the same average luminance. The target spheres were never textured to improve comparability with previous work. White lines were drawn along the edges of the box to help facilitate perception of its structure in the texture-off conditions. The lines were also shown in the texture-on conditions for consistency.

The default sphere colour was bright gray. The active target, i.e., the sphere participants were to select, was red (the top-most target in Figure 4). In the highlighting-on conditions, spheres changed blue when intersected by the stylus tip. In the highlighting-off conditions, spheres did not change colour when touched.

Head and stylus tracking used NaturalPoint's SDK. Since the stylus tip marker would make selection with a single point cursor difficult – it would occlude targets – we treated the stylus tip as a 0.25 cm radius sphere cursor instead. This cursor was not actually displayed. A hit was recorded if the participant pressed the button while the cursor sphere intersected the target sphere. In highlight-on conditions, the target sphere changed colour upon being intersected by the cursor sphere. Note that this made selection somewhat easier: targets were effectively 0.25 cm larger.

Procedure

Participants were first greeted by the experimenter who then demonstrated all conditions. After providing informed consent, participants briefly practiced (~20 to 30 selection trials) each condition. This introduction took less than five minutes. Subsequently, participants began the experiment. Participants were instructed to select the red target sphere as quickly and accurately as possible, maintaining a consistent speed/accuracy tradeoff. Selection involved intersecting the 0.25 cm "cursor" with the target sphere. This necessitated depth movement, unlike previous work, which focused on screen-space selection [37]. Hence we report the 3D size and distance of targets below, rather than adjusting these for perspective.

After each selection trial, the active target advanced according to the pattern shown in Figure 1, regardless if the trial ended in a hit or a miss. Per ISO 9241-9, the same ordering was always used to avoid influencing the task with visual search or other cognitive tasks. Upon completing all trials in a circle of targets, the next circle of targets appeared with different target size, distance, and depth values. Participants could take breaks when the top target was active (as in Figure 4), as timing began after selecting it.

At the end of the experiment, participants completed a brief questionnaire that recorded their demographic information and preferences for the conditions.

Design

The experiment used the following independent variables with the specified levels:

Texture: on or off
Highlighting: on or off
Cylinders: on or off
Target Size: 1.0 cm, 1.5 cm, 2.0 cm
Target Distance: 3.5 cm, 7.5 cm, 9.5 cm
Target Depth: 0 cm, 5 cm, 10 cm

The $2 \times 2 \times 2 = 8$ combinations of texture, highlighting, and cylinders were counterbalanced by a balanced Latin square. Target size, target distance, and target depth were selected randomly (without replacement) for each circle of 11 targets. All combinations of these factors appeared (only once) for each texture-highlight-cylinder condition. Note that target size and distance were used only to provide a range of task difficulties, and were not analyzed explicitly. Target depth is measured from the screen surface. Each circle of targets consisted of eleven targets, 10 of which were recorded. In total, participants completed $2 \times 2 \times 2 \times 3 \times 3 \times 3 \times 10 = 2160$ recorded trials. Hence, over all 16 participants, our analysis is based on 34560 selection trials.

The dependent variables were movement time (milliseconds), error rate (missed target percentage), and throughput (bits per second). We also analyzed selection deviation in the z-axis (cm).

RESULTS AND DISCUSSION

Results were analyzed with repeated measures ANOVA.

Movement Time

Movement time was measured as the time to select the target. The grand mean movement time was 972 ms (SD = 319 ms). In general, participants were able to do the task quickly at slightly under 1 second per selection. Movement times across the 8 primary conditions are summarized in Figure 5.

There was a significant main effect for highlighting on movement time ($F_{1,15} = 5.9$, $p < .05$). Highlighting actually *increased* movement time – i.e., participants were slower when highlighting was enabled. Neither texture ($F_{1,15} = 0.04$, ns) nor cylinders ($F_{1,15} = 0.07$, ns) significantly affected movement time.

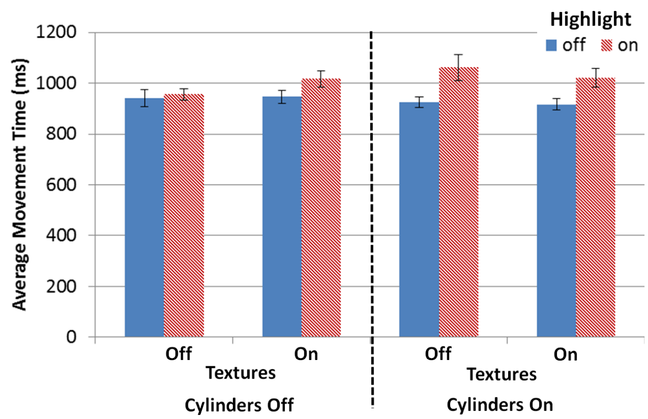

Figure 5. Movement time by textures, cylinders, and highlight. Higher scores are worse. Error bars show ±1 *SE*.

Target depth significantly affected movement time ($F_{2,15} = 8.9$, $p < .001$). The 10 cm depth condition (1151 ms) was slower than the 0 cm (863 ms) or 5 cm (904 ms) depth conditions. The interaction between highlighting and target depth was also significant ($F_{2,30}$ 6.1, $p < .01$). The slowest condition was 10 cm with highlighting-on. No other interactions were significant.

Error Rate

Error rate was the average number of selection trials that ended with a miss, expressed as a percentage of all trials for a condition. The error rates were generally high, with the best conditions hovering around 15%. Error rates are depicted in Figure 6.

There was a significant main effect for highlighting ($F_{1,15} = 29.5$, $p < .0001$). Like movement time, target depth significantly affected error rate ($F_{2,15} = 57.0$, $p < .0001$). The farther the targets were from the screen, the higher the error rate. The interaction effect between highlighting and target depth was also significant ($F_{2,30} = 15.5$, $p < .0001$). Globally, highlighting reduced the error rate by almost half – the highlighting-on conditions had 55% the error rate of the highlighting-off conditions. This was most pronounced with 10 cm targets, where error rates were as high as 42.6% with highlighting-off. This was approximately twice as high as the error rate with highlighting-on at 10 cm target depth. The lowest error rate was 11.6% for targets displayed at the screen surface (0 cm) with highlighting-on. Neither texture ($F_{1,15} = 0.38$, ns) nor cylinders ($F_{1,15} = 0.21$, ns) significantly influenced error rate. Their interactions were not significant.

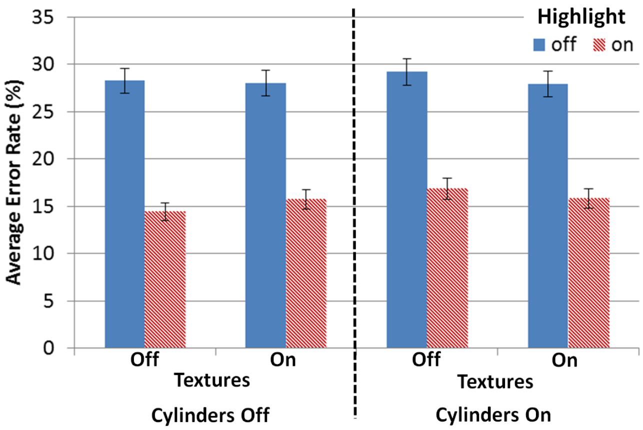

Figure 6. Error rate by textures, cylinders, and highlight. Higher scores are worse. Error bars show ±1 SE.

Throughput

Throughput was calculated with a variation we previously suggested [36] and which other researchers use [22, 36]]. The standard (2D) calculation projects the task into 1D motions along the task axis, and does not consider depth. Rather than projecting selection coordinates onto the task axis, we used the Euclidean distance between the selection coordinate and the target. The standard deviation of these distances replaced SD_x in Equation 2. Effective width and amplitude were then computed normally. This modified approach "penalizes" selections that may be visually aligned with the target, but far from it in the depth dimension. We argue that this is better suited to 3D selection. See Figure 7.

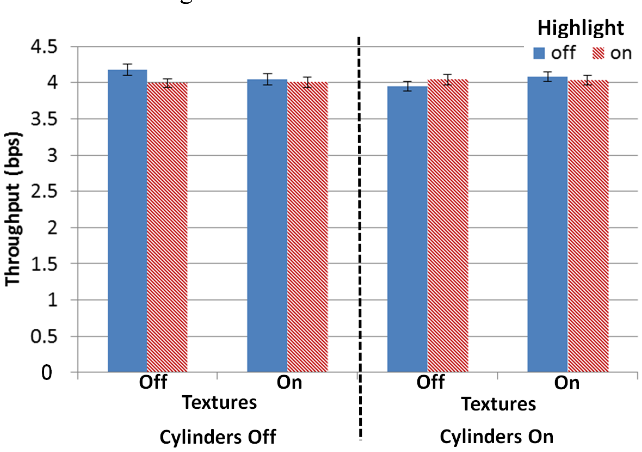

Figure 7. Throughput by textures, cylinders, and highlight. Higher scores are better. Error bars show ±1 *SE*.

Only target depth significantly affected throughput ($F_{2,15} = 31.4$, $p < .0001$) – both 0 cm and 5 cm targets had significantly higher throughput than 10 cm. Unlike movement time and error rate, highlighting was not significant ($F_{1,15} = 0.42$, ns). Texture ($F_{1,15} = 0.04$, ns) and cylinders ($F_{1,15} = 0.03$, ns) were not significant, nor were any interaction effects.

Depth Deviation

We analyzed depth deviation to address our concerns about the selection coordinate distribution. This metric was calculated as the standard deviation of selection coordinate *z* values, i.e., the "in-out" axis. Depth deviation scores are summarized in Figure 8.

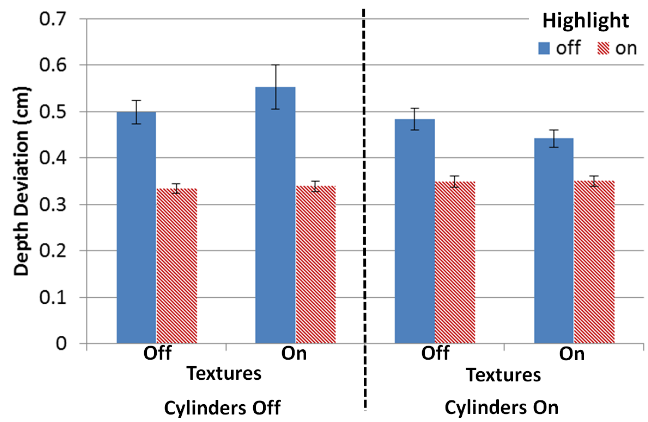

Figure 8. Depth deviation by textures, cylinders, and highlight. Higher scores are worse. Error bars show ±1 *SE*.

Like movement time and error rate, highlighting had a significant main effect on depth deviation ($F_{1,15} = 9.1$, $p < .01$). Depth deviation was significantly lower with highlighting-on, which had a mean score of 0.34 cm compared to 0.49 cm depth deviation for highlighting-off. Like the other independent variables, cylinders ($F_{1,15} = 0.56$, ns) and texture ($F_{1,15} = 0.05$, ns) had little effect, nor were their interactions significant. No interaction effects with these factors were significant either.

Target depth significantly affected depth deviation ($F_{2,15} = 23.5$, $p < .0001$). Depth deviation was lowest with 0 cm targets (0.25 cm), and was significantly higher with both 5 cm targets (0.40 cm) and 10 cm targets (0.59 cm). The interaction between highlighting and depth was significant ($F_{2,30} = 8.1$, $p < .005$). Like error rate, highlighting had a stronger effect on depth deviation for targets farther from the screen, and little effect on targets at 0 cm. This is likely due to the tactile presence of the screen, which "flattened" the selection distribution in the 0 cm condition. This makes it difficult to compare the 0 cm condition to the 5 and 10 cm conditions. Hence we primarily focus on the differences due to highlighting *within* the same depth. The effect of highlighting was strongest with the 10 cm targets: an average error of 0.73 cm with highlighting-off, vs. 0.45 cm with highlighting-on.

To facilitate comparison with previous work [9, 22], we visualized selection coordinates as scatter plots in Figure 9. The figure depicts the selection coordinates as a "side-view” for *all* cylinder-on, texture-on trials. Since this includes all 11 target locations, each coordinate has been adjusted by first transforming the target to the origin. In terms of the tracker coordinate system, the participant would sit at the left side of the plot facing right.

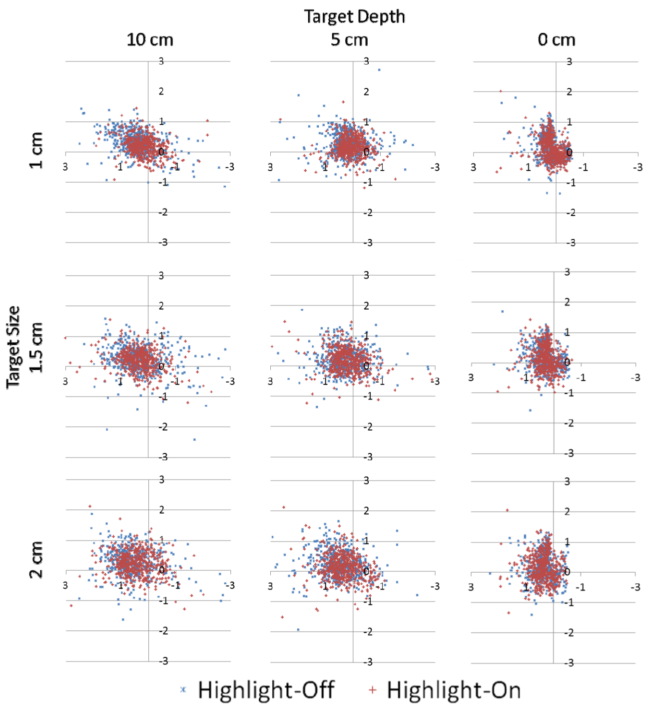

Figure 9. Selection distributions for all cylinder-on, texture-on trials, separated by target depth and target size. Highlighting-on is shown in red, and highlighting-off is shown in blue. The horizontal axis of each plot is the *z*-axis of selection coordinates, while the vertical axis shows the *y*-axis (in cm).

As seen in Figure 9, and supporting our depth deviation analysis, selection coordinates are generally more variable with greater target depths. The presence of the screen surface is visible at the 0 cm depth conditions (the right-most column), as these are relatively “flat” compared to the 5 cm and 10 cm target depths. After all, touching the screen surface stopped the stylus. Any values below 0 cm are tracking noise at the screen surface due to reflections.

The effect of highlighting is notable. The coordinates for the highlighting-off conditions are more variable, especially in the depth direction (the horizontal axis). In contrast, the highlighting-on coordinates are more circular, better approximating the target shape, as indicated by the significant effect for highlighting.

Finally, depth deviation was about 25% greater than coordinate deviation in both the x and y axes. This difference was significant ($F_{2,15} = 11.5$, $p < .0001$). However, there was significant interaction effect between the deviation axis and highlighting ($F_{2,30} = 7.2$, $p < .005$). Interestingly, this indicates that in highlighting-on conditions, depth deviation was not significantly worse than x or y axis deviation. This is reflected in the more circular patterns in Figure 9 for this condition.

Subjective Results

We asked participants about the perceived effect of each of the main independent variables. Figure 10 summarizes these results.

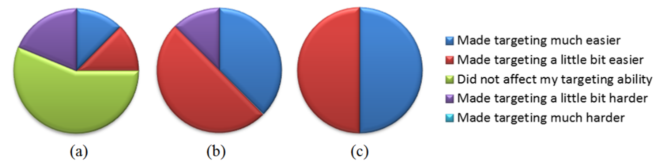

Figure 10. Participant subjective responses for (a) texturing, (b) cylinders, and (c) highlighting.

Notably, no participant reported that any of the visual aids “made targeting much harder” for any independent variable. Participants were largely ambivalent toward texture (Figure 10a) and the majority did not feel texture affected their performance. Most reported that they found cylinders (Figure 10b), and especially highlighting helpful (Figure 10c). Interestingly, the participants seemed to be aware of the relatively small effect of texturing. In contrast, they found the presence of cylinders helpful, despite the absence of quantitative results to support this.

DISCUSSION

Our results indicate that for all dependent variables investigated, texturing and the cylinders mattered very little. This is somewhat surprising, as the texture gradient on the cylinders should provide a strong depth cue [34]. Nevertheless, these factors were not significant in our analysis. While it is not in the nature of the statistical tests used to explicitly *prove* that these factors have no effect, we currently cannot reject the null hypothesis. Any effects of texturing and highlighting are certainly smaller than target depth and highlighting. Hence, we cautiously recommend that other researchers conducting 3D pointing experiments could use or ignore texturing at their discretion. However, since participants reported that they found cylinders helpful, we still recommend the use of such a “support” object for targets.

Target highlighting, on the other hand, had a significant effect on *all* dependent variables except throughput. This is in line with 2D results investigating similar effects with visually sub-optimal conditions [16]. In our case, the vergence-accommodation conflict common to stereo displays is a likely cause for the effect. This is further supported by the fact that performance (in terms of movement time, error rate, and depth deviation) was generally worse the farther targets were from the screen – stereo conflicts become more pronounced at greater depths. This is also consistent with previous work [35].

Interestingly, highlighting-on increased movement time, but decreased error rate substantially. Pointing is a classic speed-accuracy tradeoff task. Clearly, highlighting strongly influences the speed-accuracy tradeoff: participants were slower and more precise with highlighting-on. The difference in movement time was quite small, approximately 100 ms. The effect on error rate was much stronger – error rates were roughly cut in half with highlighting-on. It is likely that in the highlighting-on conditions, participants slowed down in anticipation of the appearance of the highlight. This suggests that they relied heavily on highlighting to accurately detect the 3D position of the target. On the other hand, stereo display and head-tracking alone seem to be insufficient for users to achieve the same accuracy with highlighting-off.

The high error rates (~15% in the best conditions) indicate the participants had some difficulty in the task. This is likely due to fatigue. Although the task was somewhat similar to stylus use on a tablet, our display was upright and required holding the arm up to the display. After numerous trials, this becomes fatiguing, and several participants complained of this during the experiment. It is also possible that participants of our experiment were simply inherently fast and thus inaccurate. The movement times reported here are roughly 25% lower than those in previous work [35] and error rates are roughly double. While the previous work used a similar stylus-based technique, the display was tipped on its back (instead of upright), which is likely less fatiguing. We thus believe these two factors together largely explain the high error rates.

That throughput was not influenced by highlighting reveals an important quality of the metric. Previous work pointing indicated that throughput is constant despite speed-accuracy tradeoffs [24]. Unlike MacKenzie and Isokoski, we did not explicitly ask participants to focus on speed or accuracy. Highlighting seemed to implicitly provide these emphases, however. Consequently, our work appears to be the first (indirect) confirmation of the immunity of throughput to speed-accuracy tradeoffs in 3D selection tasks. The throughput scores were also roughly consistent with previous work [35], further supporting use of the metric.

Finally, it is notable that participants were able to perform the task so well as to create a roughly spherical selection distribution (see Figure 9). Consistent with previous work [22, 40], depth deviation was higher than either the *x* or *y* axes. However, the magnitude of our depth deviations is smaller than in other recent work using head-mounted displays [22]. Perhaps most surprising is how strong the effect of highlighting was. In our experiment, highlighting substantially reduced depth deviation, yielding roughly spherical selection distributions. Lubos et al. [22] also used similar feedback, yet had substantially more ellipsoid shaped selection distributions (in the depth direction). This may be due to differences in the display technology used.

CONCLUSIONS

We presented a study investigating several issues in the visual presentation of 3D point-selection experiments. The study used the ISO 9241-9 method, and compared the effects of texturing, target "support" cylinders, and target highlighting as visual feedback. Highlighting had the strongest effect overall. With highlighting, selection time significantly increased while accuracy improved, cutting error rates almost in half. Participants also strongly preferred highlighting. Consequently, we recommend researchers consider the use of target highlighting in 3D pointing experiments. At the very least, its presence or absence should be reported to ensure comparability with other work.

Notably, throughput was largely unaffected by the influence of highlighting. As previously demonstrated [24], this is likely because throughput is largely immune to speed-accuracy tradeoffs. The effect of highlighting "skewed" this tradeoff in favour of accuracy in our experiment, yet throughput compensated for this skew. This further emphasizes the utility of the throughput measure.

Finally, in terms of recommendations to system designers, it is evident that some form of visual feedback, such as highlighting, is important in improving the usability of 3D selection interfaces. We argue that the small cost in terms of movement time is acceptable considering the large improvement in error rates. Highlighting may be somewhat distracting in cluttered environments, but this is a topic for future work. Our other visual aids, texturing, and using cylinders to prevent objects from floating had relatively little effect. Nevertheless, we suggest that researchers conducting 3D pointing experiments should be clear to report these design decisions as they may affect the comparability of their results.

While we focused exclusively on touch-based 3D point selection, in future work, we will investigate highlighting, cylinders, and texture on remote pointing techniques.

Acknowledgements

Thanks to Eduardo Soto for assistance running the experiment. This work was supported by NSERC.

REFERENCES

1. Achibet, M., Marchal, M., Argelaguet, F., and Lécuyer, A., The Virtual Mitten: A novel interaction paradigm for visuo-haptic manipulation of objects using grip force, *Proc. of the IEEE Symposium on 3D User Interfaces - 3DUI 2014*, (New York: IEEE, 2014), 59-66.
2. Apitz, G., Guimbretière, F., and Zhai, S., Foundations for designing and evaluating user interfaces based on the crossing paradigm, *ACM Transactions on Computer-Human Interaction*, 17, 2010, 9.
3. Argelaguet, F. and Andujar, C., A survey of 3D object selection techniques for virtual environments, *Computers & Graphics*, 37, 2013, 121-136.
4. Boritz, J. and Booth, K. S., A study of interactive 3D point location in a computer simulated virtual environment, *Proc. of the ACM Symposium on Virtual*

Reality Software and Technology - VRST '97, (New York: ACM, 1997), 181-187.

5. Boritz, J. and Booth, K. S., A Study of Interactive 6 DOF Docking in a Computerized Virtual Environment, *Proc. of the Virtual Reality Annual International Symposium*, (New York: IEEE, 1998), 139-146.
6. Bowman, D. A, Johnson, D. B., and Hodges, L. F., Testbed evaluation of virtual environment interaction techniques, *Proc. of the ACM Symposium on Virtual Reality Software and Technology - VRST '99*, (New York: ACM, 1999), 26-33.
7. Bowman, D. A. , Kruijff, E., LaViola, J. J. , and Poupyrev, I., 3D User Interfaces: Theory and Practice. Addison Wesley Longman Publishing Co., Inc., 2004.
8. Bowman, D. A., Wingrave, C. A., Campbell, J. M., Ly, V. Q., and Rhoton, C. J., Novel uses of pinch gloves for virtual environment interaction techniques, *Virtual Reality*, 6, 2002, 122 - 129.
9. Bruder, G., Steinicke, F., and Stuerzlinger, W., Touching the void revisited: Analyses of touch behavior on and above tabletop surfaces, *Proc. Human-Computer Interaction–INTERACT 2013*. Springer, 2013, 278-296.
10. Bruder, G., Steinicke, F., and Sturzlinger, W., Effects of visual conflicts on 3D selection task performance in stereoscopic display environments, *Proc. of the IEEE Symposium on 3D User Interfaces - 3DUI 2013*, (New York: IEEE, 2013), 115-118.
11. 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.
12. Chun, K., Verplank, B., Barbagli, F., and Salisbury, K., Evaluating haptics and 3D stereo displays using Fitts' law, *Proc. of the IEEE International Workshop on Haptic, Audio and Visual Environments and Their Applications*, (New York: IEEE, 2004), 53-58.
13. Cockburn, A., Quinn, P., Gutwin, C., Ramos, G., and Looser, J., Air pointing: Design and evaluation of spatial target acquisition with and without visual feedback, *International Journal of Human-Computer Studies*, 69, 2011, 401-414.
14. Dang, N.-T., A survey and classification of 3D pointing techniques, *Proc. of the IEEE International Conference on Research, Innovation and Vision for the Future*, (New York: IEEE, 2007), 71 - 80.
15. 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.
16. Fraser, J. and Gutwin, C., The effects of feedback on targeting performance in visually stressed conditions, *Proc. of Graphics Interface*, (Toronto: Canadian Human-Computer Communication Society, 2000).
17. 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.
18. Hubona, G. S., Wheeler, P. N., Shirah, G. W., and Brandt, M., The relative contributions of stereo, lighting, and background scenes in promoting 3D depth visualization, *ACM Transactions on Computer-Human Interaction*, 6, 1999, 214-242.
19. 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.
20. 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.
21. Liu, L., Liere, R. v., Nieuwenhuizen, C., and Martens, J.-B., Comparing aimed movements in the real world and in virtual reality, *Proc. of the IEEE Virtual Reality Conference - VR 2009*, (New York: IEEE, 2009), 219-222.
22. Lubos, P., Bruder, G., and Steinicke, F., Analysis of direct selection in head-mounted display environments, *Proc. of the IEEE Symposium on 3D User Interfaces - 3DUI 2014*, (New York: IEEE, 2014), 11-18.
23. MacKenzie, I. S., Fitts' law as a research and design tool in human-computer interaction, *Human-Computer Interaction*, 7, 1992, 91-139.
24. MacKenzie, I. S. and Isokoski, P., Fitts' throughput and the speed-accuracy tradeoff, *Proc. of the ACM SIGCHI Conference on Human Factors in Computing Systems - CHI 2008*, (New York: ACM, 2008), 1633-1636.
25. Mould, D. and Gutwin, C., The effects of feedback on targeting with multiple moving targets, *Proc. of Graphics Interface 2004*, (Toronto: Canadian Human-Computer Communications Society, 2004), 25-32.
26. Nieuwenhuizen, K., Liu, L., Liere, R. v., and Martens, J.-B., Insights from Dividing 3D Goal-Directed Movements into Meaningful Phases, *IEEE Computer Graphics and Applications*, 29, 2009, 44 - 53.
27. Pawar, V. M. and Steed, A., Profiling the behaviour of 3D selection tasks on movement time when using natural haptic pointing gestures, *Proc. of the ACM Symposium on Virtual Reality Software and Technology - VRST 2009*, (New York: ACM, 2009), 79-82.

28. Poupyrev, I., Ichikawa T., Weghorst, S., and Billinghurst, M., Egocentric object manipulation in virtual environments: Empirical evaluation of interaction techniques, *Proc. of Eurographics '98*, (New York: ACM, 1998), 41-52.

29. Prachyabrued, M. and Borst, C. W., Visual feedback for virtual grasping, *Proc. of the IEEE Symposium on 3D User Interfaces - 3DUI 2014*, (New York: IEEE, 2014), 19-26.

30. 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.

31. 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.

32. Sprague, D. W., Po, B. A., and Booth, K. S., The importance of accurate VR head registration on skilled motor performance, *Proc. of Graphics Interface 2006*, (Toronto: Canadian Information Processing Society, 2006), 131-137.

33. Stuerzlinger, W., Considerations for targets in 3D pointing experiments, *Workshop on Interactive Surfaces for Interaction with Stereoscopic 3D - ISIS3D 2013.*

34. Surdick, R. T., Davis, E. T., King, R. A., Corso, G. M., Shapiro, A., Hodges, L., and Elliot, K., Relevant cues for the visual perception of depth: is where you see it where it is?, *Proc. of the Human Factors and Ergonomics Society Annual Meeting*, (SAGE Publications, 1994), 1305-1309.

35. Teather, R. J., Pavlovych, A., Stuerzlinger, W., and MacKenzie, I. S., Effects of tracking technology, latency, and spatial jitter on object movement, *Proc. of the IEEE Symposium on 3D User Interfaces - 3DUI 2009*, (New York: IEEE, 2009), 43-50.

36. 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.

37. 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.

38. Valkov, D., Steinicke, F., Bruder, G., and Hinrichs, K., 2D touching of 3D stereoscopic objects, *Proc. of the ACM Conference on Human factors in Computing Systems - CHI 2011*, (New York: ACM, 2011), 1353-1362.

39. Zhai, S., Buxton, W., and Milgram, P., The “Silk Cursor”: Investigating transparency for 3D target acquisition, *Proc. of the ACM Conference on Human Factors in Computing Systems - CHI '94*, (New York: ACM, 1994), 459 - 464.

40. Zhai, S., Milgram, P., and Rastogi, A., Anisotropic human performance in six degree-of-freedom tracking: An evaluation of three-dimensional display and control interfaces, *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, 27, 1997, 518-528.