Assessing the Effects of Orientation and Device on (Constrained) 3D Movement Techniques

Robert J. Teather *

Wolfgang Stuerzlinger †

Department of Computer Science & Engineering,
York University, Toronto

ABSTRACT

We present two studies to assess which physical factors influence 3D object movement tasks with various input devices. Since past research has shown that a mouse with suitable mapping techniques can serve as a good input device for *some* 3D object movement tasks, we also evaluate which characteristics of the mouse sustain its success.

Our first study evaluates the effect of a supporting surface across orientation of input device movement and display orientation. A 3D tracking device was used in all conditions for consistency. The results of this study are inconclusive; no significant differences were found between the factors examined. The results of a second study show that the mouse outperforms the tracker for speed in all instances. The presence of support also improved accuracy when tracker movement is limited to 2D operation. A 3DOF movement mode performed worst overall.

CR Categories and Subject Descriptors: H.5.1 [**Information Interfaces and Presentation**]: Multimedia Information Systems – *virtual reality.* H.5.2 [**Information Interfaces and Presentation**]: User Interfaces – *input devices, interaction style.*

Additional Keywords: 3D manipulation, comparing devices

1 INTRODUCTION

Many research studies have targeted the development of intuitive 3D manipulation techniques for virtual environments. However, to this day, it is still far more difficult to perform simple tasks in a virtual reality (VR) setup compared to conceptually similar tasks in a desktop environment. Consider, for example, the relative ease of moving a desktop icon, and then compare this to the problem of moving an object in a 3D virtual environment.

Most previous research focuses on creating better 3D manipulation techniques for use with 3D input devices such as trackers and wands, which allow the user to control up to 6 degrees of freedom (DOFs) simultaneously. However, the mouse often outperforms these devices for common tasks in many systems, although 3D devices seem better suited to the task. User familiarity may play a big factor here; most people use a mouse extensively in day-to-day computing and have very limited experience with 3D devices. Another factor is the dimensionality of the task. It is more difficult to accurately position an object in 3D space than in 2D space, mainly due to the additional degree(s) of freedom in which the object can move. Another factor is that the mouse requires a supporting surface on which to operate. This supporting surface reduces fatigue and hand jitter of the user, providing an advantage over the "free-floating" movement associated with most 6DOF devices. On the other hand, this is also a disadvantage for the mouse, as it is then unsuitable for virtual environments that require full 6DOF movement or for VR setups where a supporting surface is impractical (e.g. CAVEs).

Furthermore, many VR input techniques couple the display space to the input space, and register the position of virtual objects or cursors with the user's real hand(s). Conversely, the mouse is an indirect, relative manipulation device, which is decoupled from display space. In addition, the mouse moves in a horizontal movement plane, which is mapped to a vertical movement plane on a typical desktop computer.

We aim to evaluate to what extent these physical factors – display orientation, input device movement orientation, physical support, and device characteristics – affect 3D movement tasks. In particular, we evaluate the effect of a supporting surface, as required for a mouse. Orientation of the display relative to the input device's movement is also considered. This is to determine the differences between a direct mapping (e.g. device movement up to cursor movement up) and the indirect mapping used by the mouse (e.g. device movement forward to cursor movement up).

The overall goal of this work is to investigate why the mouse is so well-suited to certain types of constrained 3D movement tasks. A secondary goal is to determine how these factors can also benefit the design of movement techniques with 3D input devices.

* email: rteather@cse.yorku.ca

† www: http://www.cse.yorku.ca/~wolfgang

2 RELATED WORK

Previous work on 3D manipulation, especially with 2D input devices, and the use of supporting surfaces is examined below.

2.1 3D Manipulation

A large variety of previous work addresses the use of 6DOF input devices for 3D manipulation tasks [3, 4, 5, 13, 16]. A general 3D manipulation task includes both positioning and rotation, and requires selection of the object to be manipulated prior to manipulation. Selection is accomplished either through the use of a 3D cursor/hand for direct selection or ray casting.

Ray casting has been found to be an excellent selection technique for 3D devices [5, 16, 22] and is commonly used in VR systems. Once an object has been selected, its 3D position is then linked to the 3D position of the 6DOF device. Moreover, ray casting also enables 3D selection with 2D input devices. For this the mouse cursor position on the display is used to generate a ray from the viewpoint through that 2D point into the scene. Once the first object hit by that ray is selected, software techniques are required to map 2D mouse motions into 3D movement operations. The majority of such mappings require the user to mentally decompose tasks into a series of 1 or 2DOF operations along the coordinate system axes. Examples for this are 3D widgets, such as "3D handles" [7] and the "skitters and jacks" technique [2], or modes, such as those used by the "DO-IT" technique [13].

However, some systems also support 3D direct manipulation similar to the drag 'n' drop paradigm prevalent in desktop GUIs. Designers of these systems typically make a set of assumptions, which permit users to leverage their familiarity with 2D desktop environments into the domain of 3D virtual environments. All of these techniques introduce some kind of constraints to achieve this. At the simplest level, gravity and collision avoidance are used to ensure objects rest on the ground, and do not interpenetrate each other. A more advanced approach involves pre-programming specific constraints such that objects behave according to human expectations. For example, a desk rests on the floor; a desk lamp sits on top of the desk, etc. [17].

Recent work introduced a more generalized sliding paradigm in which objects always stick to other objects in the scene, and slide along their surfaces when dragged with the mouse. This uses the constraint that (almost) all objects in the real world are attached to other objects. This sliding technique was empirically demonstrated to be superior to indirect approaches such as 3D widgets [14], and was also shown to outperform 3DOF movement techniques for certain types of scene assembly tasks [19].

2.2 Physical Support and Passive Haptic Feedback

The mouse requires a physical surface upon which to operate. This is both an advantage and a limitation of the device. It helps prevent fatigue as users can rest their arm and also prevents jitter that can decrease accuracy of object movements. However, it also renders the mouse largely unsuitable for certain VR setups such as CAVEs, since it constrains usage to locations where a tabletop or similar surface is present. This problem is exacerbated in virtual environments using head-mounted displays, as the user is also unable to see the mouse itself [10].

Nevertheless, the benefits of support have not gone unnoticed in the VR and AR communities. Previous work attempted to combine the best of both worlds by adding a mobile physical support surface to traditional VR setups. Most notable among these are the HARP system [10], the Virtual Notepad [15] and the Personal Interaction Panel [18]. These approaches present virtual interfaces overlaid over a real physical surface (often called a slate or paddle), which the user carries with them. Other work used the non-dominant hand directly for support [9]. The virtual representation of the slate can feature either 2D or 3D widgets. The goal of these interfaces is to leverage the best aspects of 2D *and* 3D user interfaces, i.e. a 3D virtual environment, in which the user can navigate, coupled with a more familiar 2D interface. Typically a 6DOF input device (e.g. a tracked stylus) is used to determine if the user interacts with the slate and which UI widgets are being selected. An alternative is to utilize a secondary input device, such as a tablet PC, as the slate [6].

Other work has compared 3D interaction on and off tabletop surfaces, to assess the importance of passive haptic feedback in a display/input coupled environment [20]. They found that object positioning was significantly faster due to the support offered by the tabletop surface, but that accuracy was slightly worse.

3 Comparing Input Devices

Our goal is to determine the relative importance of various factors that distinguish 3D interaction with a mouse from interaction with 6DOF input devices. Thus, we chose to compare interaction with and without a supporting surface, as well as the effects of input device movement orientation and display orientation. However, directly comparing two different input devices is problematic since it can be extremely difficult to account for all possible confounding factors that affect their performance.

One potentially confounding factor is clearly any differences in control space orientation [23]. Another is different hand positions used with different input devices. Both of these factors also relate to specific muscle groups that may be more or less developed and can affect fine motor control [24]. In particular, input devices that use fine-motor control muscle groups, such as those in the fingers, can benefit precision manipulation. However, allowing several muscle groups in the arm to work together, rather than in isolation can be even better. This is supported by later work comparing muscle groups in the fingers, wrist and forearm. The results show that using multiple muscle groups together tended to perform better than just using the fingers alone [1]. Technical properties such as tracking accuracy and jitter levels can also impact performance. Furthermore, large differences in movement distances and/or cursor speed may also play a role.

Consequently, we designed our test environment to eliminate as many of these factors as possible. One of the main decisions for our first study was to use a 3D tracker as the input device for all conditions. However, we also required the user to hold a mouse in the palm of their hand. This “flying mouse” device combination is very similar to the Bat [21].

To evaluate the supporting surface while keeping the input device constant, we chose to have users move the tracker/mouse on a table. This effectively uses the tracker to emulate a mouse. However, the devices are not identical, as the mouse permits “clutching”, i.e., picking up the device to reposition it for long distance movements. As a 3D tracker is an absolute positioning device, we used a direct mapping between a rectangular region on (or off) the supporting surface and the display. Thus, the tracker behaves similarly to a graphical tablet or “puck”, i.e., device position in a rectangular region maps directly to screen position.

3.1 General Assumptions about 3D Manipulation

While designing our studies, we made several assumptions about 3D positioning. These assumptions are based on empirical results, and conform to generally accepted 3D UI design practices.

The first assumption is that ray casting is a better choice than direct 3D selection with 3D devices [5, 16]. Other work indicates that ray-casting is also well-suited for 2D devices, and even outperforms 3D devices [22] for selecting 3D objects.

A second assumption is that objects can be constrained to remain in contact with the remainder of the scene at (almost) all times [14, 19]. This is based on the observation that in the real world, gravity ensures that objects do not float in space. Hence, contact is the *appropriate* default for most virtual environments, with the exception of flight and space simulations. Experiments revealed that the contact assumption is particularly beneficial for novice users, but even experts profit from it [8, 14].

A third and final assumption is that collision avoidance benefits 3D manipulation. Fine positioning of objects is greatly aided by the ability to slide objects into place with collision avoidance [8]. One reason for the effectiveness of collision avoidance is that novice users of graphical systems often become confused when objects interpenetrate one another and experience difficulties in resolving the problem. After all, solid objects in the real world never interpenetrate. Hence, this is the *proper* default [8, 14, 19].

We believe that these design decisions greatly improve the immediate usability of VR systems – which otherwise can require a great amount of training and are then only usable by experts.

3.2 3D Movement Technique

The 3D movement technique used here relies on the idea of contact-based sliding. It is based on the “contact” assumption discussed above and uses ray casting for selection. The sliding technique ensures that the object being moved remains stably “under” the cursor, yet in contact with other objects in the scene at all times [14]. Depth is handled automatically; objects simply slide across the closest surface relative to the viewer that their projection falls onto. This effectively reduces 3D positioning to a

2D problem, as objects can now be directly manipulated via their 2D projection. It also makes 3D manipulation similar to drag 'n' drop interfaces in modern desktop computing, except that it also affects the 3D position of objects. We chose this technique because user studies have shown that novices find it much easier to use compared to other VR techniques, such as 3D widgets [14]. We were also interested in determining how well this technique can be used with 3D input devices and how it compares.

4 EXPERIMENTS

Two user studies were conducted to empirically evaluate the relative importance of the factors discussed above.

4.1 First Study: Support and Orientation

This study compared the main factors being examined: hand support, display orientation and device movement plane orientation. The goal was to determine to what extent physical support aids the mouse in constrained 3D movement tasks. A secondary goal was to determine if matching the input device movement orientation to the screen orientation resulted in better performance than mismatched situations.

4.1.1 Hypotheses

Based on the results of previous work, we hypothesized that participants would perform better overall in the supported conditions. The physical surface allows the user to rest their arm and hence reduces hand jitter, improving accuracy. Due to the inherent speed/accuracy trade-off in this type of object movement task, we predicted that speed would also improve, as they would have to spend less time trying to accurately position objects.

We also hypothesized that the standard desktop display/device orientation combination would prove to be the best, due to the participants' familiarity with it. However, we also believed that users would generally perform better in conditions in which the movement plane of the input device matched that of the display, due to the direct mapping of input motions to cursor movement.


a

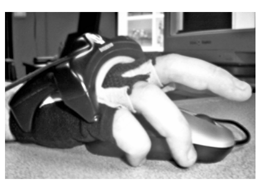
b

Figure 1. a) The experimental setup. The table to the right of the displays was used for the horizontal support condition, and the cupboard resting on top for the vertical support condition. The whole table was removed in the no support condition. b) Hand tracker and mouse – two fingers lifted to show mouse underneath.

4.1.2 Participants

Sixteen paid participants took part in the study. Their ages ranged from 18 to 28, with a mean of 22.45 years. Only one participant was female. The average mouse usage for the group was 11.9 years. All participants used the mouse with their right hand.

4.1.3 Apparatus

Tasks were performed in a desktop VR system (Figure 1a), consisting of a desktop PC with stereoscopic graphics and 3D input. This was an Intel Pentium 4 at 3GHz with 512MB RAM, and an NVidia Quadro FX3400 graphics card. Two SGI monitors with 800 x 600 at 120 Hz were used for stereo display. Brightness and colour of these displays was adjusted to be as similar as possible. One monitor was positioned upright, and the other was supported on its back with hard Styrofoam. The horizontal monitor was inclined ~10° for more ergonomic viewing, while still maintaining approximate orthogonality to the vertical monitor. LCD shutter glasses and a Stereographics emitter were used for stereo viewing. Room lights were dimmed to equalize glare across both displays, since this could affect stereo viewing.

An Intersense IS900 was used for tracking the 3D position of the user's right hand. In this hand, participants also held an optical mouse and its buttons were used to record "click" events. The optical sensor of the mouse was taped over. All cursor/object movement was recorded only by the 3D tracker, which was mounted on the back of a nylon glove worn in all conditions. Figure 1b depicts the position of tracker and mouse on a hand.

Since the tracker is an absolute positioning device, a small rectangle (15x11.25 cm) was marked out on the table, to visualize the mapping of movement to cursor movement on the screen. This area has the same height/width ratio as the screen. Upon starting each trial, the software registered the position of the tracker as the bottom left corner of the screen, and placed the cursor there. Participants were required to place their hand in that position at the start of each trial. Hand support was provided by a table in the horizontal device movement condition, and a sturdy cupboard on top of the table for the vertical input device movement condition. These were moved out of the way in the unsupported conditions. Small marks on the floor and tabletop ensured that the physical supports were always in the same position when in use.

The software was written in C++ with OpenGL and included stereo pair rendering to generate the stereoscopic graphics effect. It used the sliding movement technique described in section 3.2.

4.1.4 Procedure

After an introduction and signing informed consent forms, each participant was seated in front of the system and given the shutter glasses and tracked glove to wear. They were then given a single practice trial to familiarize them with the task.

The experimental task (Figure 2) involved moving several pieces of furniture around a computer lab virtual environment. Participants were initially presented with a low-angle view of the scene, similar to Figure 2a.

The task required that they move two computer stations to foreground desks, as well as a chair. A printer had to be moved from the second row to the back-most desk, and a stack of books from the front-most desk to the second row, right-most desk. Overall, the task involved moving object *1* to position *A*, object *2* to position *B*, and so on, as depicted in Figure 2b. Figure 2d shows the completed scene from an overhead view. Although complex, the task was intended to assess performance in a fairly realistic scenario, rather than examine abstract motions. This task was selected to make the results more generalizable.

Moving a computer station involved moving both the monitor and the keyboard. Users were not required to move the mouse objects in the model, because a pilot study found that it was too small to be selected reliably in some of the conditions. Thus the mouse object was excluded to ensure that the task could be completed under all conditions. In total, each trial involved the movement of 7 virtual objects, of sizes ranging from relatively small (the books) to relatively large (monitor and printer).

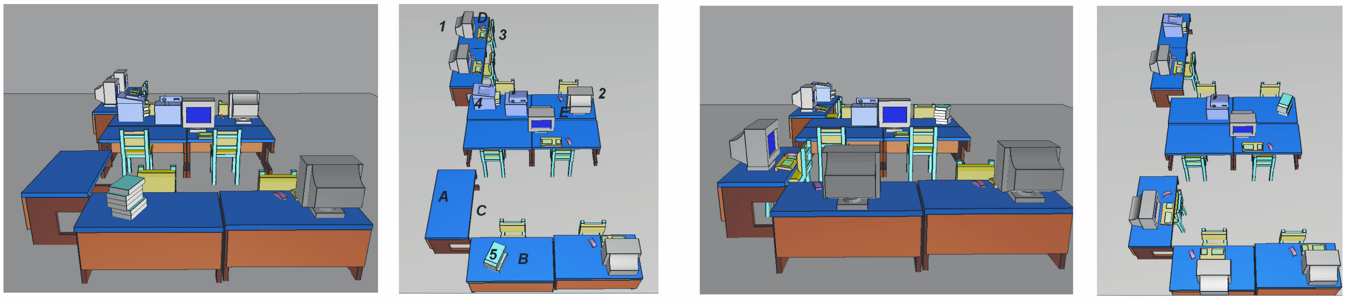

Figure 2. a) View of starting condition (this is what the participants saw for the first study), b) Overhead view of starting condition (for illustration only), c) View of target scene, d) Overhead view of target scene.

A certain degree of selection accuracy was also required in this task. For example, selecting the top book in the stack would only move that book; participants had to select the bottom book to move the entire stack.

Participants were given continuous verbal feedback throughout the experiment as well as reminders on the ordering if they showed signs of confusion about which object to move next. After two or three repetitions, they were usually able to remember the sequence without aid from the experimenter. Scene rotation was enabled, and participants were allowed to change the viewpoint (accomplished via a drag on the background of the scene). However, participants were encouraged to use a top-down view, similar to Figure 2b, as it made the task easier. Virtually all of them changed the viewpoint to this perspective in each trial.

Participants were also encouraged to take breaks between trials, particularly in the vertical device conditions, as these were the least ergonomic and most fatiguing. A counterbalanced ordering also helped ensure that participants did not spend extended periods of time in these conditions. Following the experiment, they were surveyed for subjective preferences as well.

4.1.5 Design

The experiment was a 2×2×2×4 within-subjects design. The independent variables were display orientation (vertical and horizontal) input device movement orientation (vertical and horizontal), support (supported or unsupported) and trial number (1 through 4), respectively. Figure 3 depicts all 8 combinations of the independent variables.

The orderings of support and device orientation were counterbalanced according to a balanced Latin square to compensate for learning effects across conditions. To reduce the effect of the relatively large time required to switch the display between the top to bottom monitor, half of the participants completed all trials in the vertical display condition first, followed by the horizontal display condition. The other half used the horizontal display first followed by the vertical.

Participants performed the task a total of 32 times. Overall, it took approximately 1 hour to complete the series of trials.

4.1.6 Results

The dependent variables were task completion time and accuracy. Accuracy was measured by summing the straight-line distances between object positions at the end of the task compared to the target scene. Mean task completion times and accuracy measures with standard deviations are shown in Figures 4 and 5.

A repeated measures ANOVA found no significant main effect on completion time for display orientation ($F_{1,511}$=0.25, ns), device movement orientation ($F_{1,511}$=0.48, ns), or hand support ($F_{1,511}$=0.05, ns). A significant effect for trial number ($F_{3,511}$=8.07, $p<.05$) was found, indicating that participants got faster with practice. An interaction between trial number and device orientation fell just short of significance ($F_{3,511}$=2.73, p=.055).

For another analysis we split all trials into two groups: one where input device movement orientation and display orientation matched, and one where they did not. There was no significant difference ($F_{1,511}$=0.02, ns). We also compared the effect of display orientation ordering. Participants who first completed the vertical display and then the horizontal, had a mean completion time of 65.52s and were significantly faster than the 67.24s for participants who did the horizontal display first ($F_{1,511}$=5.06, $p<.05$). However, if the first trial from each condition is excluded, this difference was not significant ($F_{1,383}$=2.26, $p>.05$).

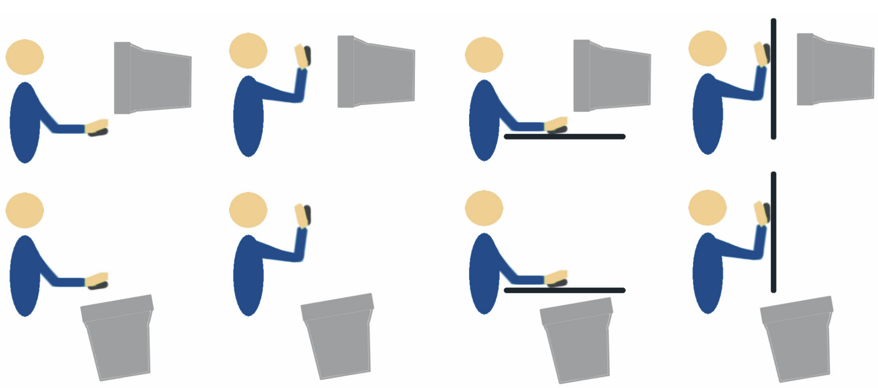

Figure 3. The eight experimental conditions. The left four represent the unsupported conditions, and the right four the supported conditions. The top four represent the vertical display, and the bottom four represent the horizontal display.

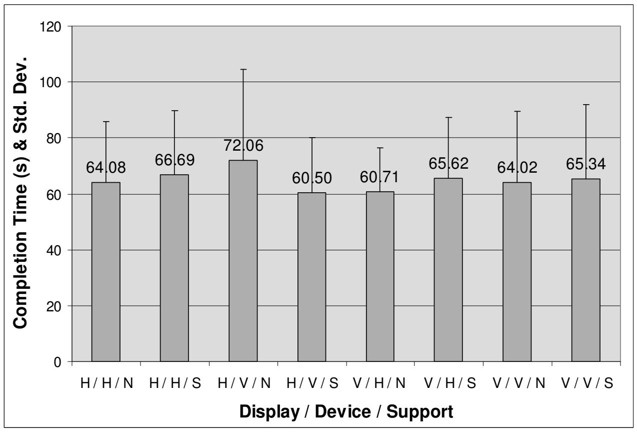

Figure 4. Mean task completion times by condition (study 1).

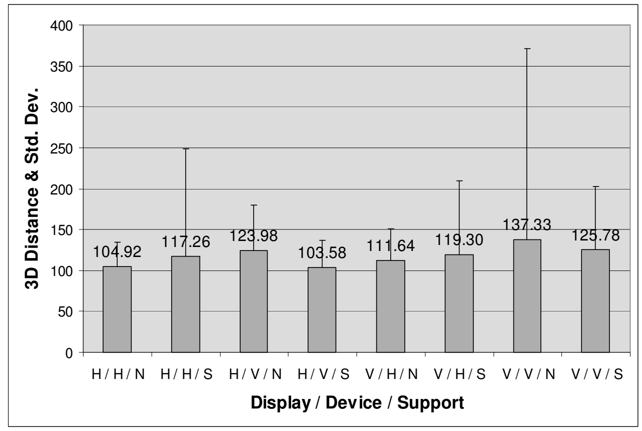

Figure 5. Mean error distance by condition (study 1).

Due to a software logging error, one accuracy log file was lost. Thus, only 511 such measures were recorded. For accuracy, no significant difference was found in the three conditions: display orientation ($F_{1,510}$=0.95, ns), device orientation ($F_{1,510}$=1.44, p>.05) and support ($F_{1,510}$=0.17, ns). No significant effect for display ordering was found on accuracy ($F_{1,510}$=0.44, ns).

Fourteen of the sixteen participants replied to the questionnaire. Of these responses, half preferred support, and half did not. The display/device orientation combinations were ranked in order of preference on a scale of 1 to 4, with 1 being most preferred. The ranks for these combinations were analyzed with a Kruskal-Wallis ANOVA and were found to be significantly different (H_3=26.32, p<0.0001). The mean rankings for each combination were 1.42 for the "standard desktop" (vertical display, horizontal device = "VH") configuration, 2.14 for the "HH" condition, 2.86 for the "VV" condition, and 3.57 for the "HV" configuration.

4.1.7 Discussion of Device and Display Orientation

The results of the first study are inconclusive and we could not determine if input device orientation and display orientation affect performance in constrained 3D movement tasks. Moreover, the statistical power of all tests was fairly low (in the range 0.1–0.2), suggesting that many more participants would be required to *reliably* detect significant results for the conditions. The maximum difference between similar conditions is also less than 20%, i.e. the magnitude of any potential effect is also limited. Only the nearly significant interaction between trial and device orientation shows that participants were almost significantly better with the horizontal device condition by the fourth repetition compared to vertical. Considering that significant improvements were observed with practice, it seems likely that this interaction effect could become significant with additional repetitions. However, it is not surprising that users might get better faster with the horizontal device; not only is this condition more ergonomic but it is also more familiar due to its similarity to the mouse.

During the experiment, we observed that participants often moved the device diagonally in the unsupported conditions. This was impossible in the supported conditions, as the supporting surfaces physically prevented it – device movement was constrained to either the vertical or horizontal 2D plane. This could explain why no significant effect was found for device orientation. However, if motion was diagonal in all unsupported conditions, we could expect asymmetric learning: users should get better faster in the unsupported conditions. However, no evidence of this was found. This may suggest that proprioception alone is insufficient for users to accurately move in a single plane of motion in free space. Several participants' comments support this: they were able to constrain their hand motion to the 2D plane *if* they watched their hand, but not when relying solely on proprioception (i.e. without looking at their hand).

Since display ordering showed an effect on task completion times, it seems that counterbalancing was not completely successful. However, the effect was quite small (about 2% difference) and disappears when the first trial from each condition is excluded (i.e. the difference disappears with practice). In addition, nothing is evident in terms of accuracy. Thus we attribute this to the relative unfamiliarity of a horizontal display.

One potential confound in this study is that participants were allowed to freely rotate the scene. However, observations during the experiment show that the scene rotation itself took only about 1–2 seconds (i.e. a very small percentage of the overall time). Moreover, virtually every participant rotated to (nearly) the same overhead view in each trial.

Overall, the lack of significant effects prompted the design of our second study. We decided to focus on the support condition. Consequently, all other factors where no significant differences were found were "collapsed" and only the vertical display and the horizontal device movement conditions were used in the second study. This was done to decrease the variability between conditions and to focus on any potentially significant effects.

4.2 Second study: Mouse and 3D Tracker

The goals of this study were to further evaluate physical support, and to determine what other features of the mouse make it a good input device for constrained 3D positioning. Consequently, we decided to directly compare the mouse to the 3D tracker in several conditions, including the 2D movement modes used above as well as a full (i.e. unconstrained) 3DOF movement mode.

4.2.1 Hypotheses

The first hypothesis of this study was that the mouse would outperform the tracker in all conditions. This could indicate that the most plausible explanation for the results of the first study is one of the features that were not investigated in that study. One such feature is tracking resolution. Based on previous work [14, 19], we also predicted that an unconstrained 3DOF tracker would be slower than all other conditions, including the 2D constrained tracker conditions.

4.2.2 Participants

Ten paid participants took part in the study. Ages ranged from 19 to 26 years, with a mean of 22.1 years. Five were male, and five were female. All used the computer mouse with their right hand, with an average of 13.4 years of usage.

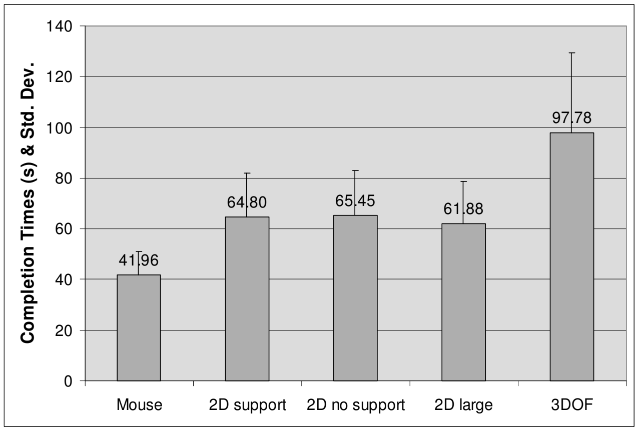

Figure 6. Mean task completion times by condition (study 2).

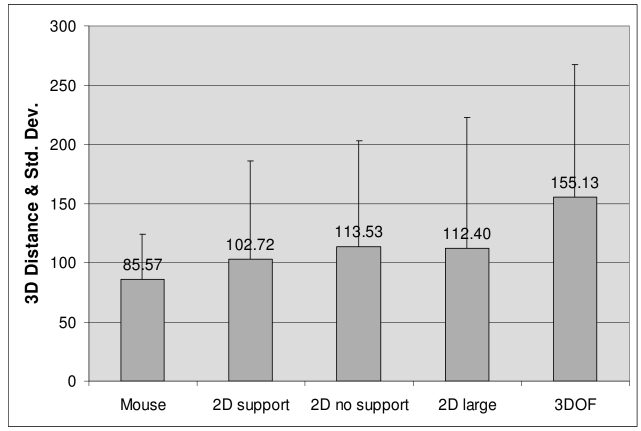

Figure 7. Mean error distances by condition (study 2).

4.2.3 Apparatus

Tasks were performed in the same desktop VR system, using the same displays and stereoscopic system. Only the vertical monitor was used in this study.

This study used an optical mouse as well as the IS900 tracker used in the first one. One of the five conditions used the mouse, with its speed set to match the tracker as closely as possible, and all acceleration/enhancements disabled. All other conditions used the 3DOF tracker in a variety of modes. The tracker was worn in all conditions. The table was used to support the mouse and the supported tracker conditions. The tracker again operated as an absolute positioning device. Most of the tracker conditions used the same 15x11.25 cm rectangle to represent the mapping to the screen. However, one condition increased the area to 30x22.5 cm to investigate the effect of an increased relative tracking resolution. This mode provided approximately a one-to-one correspondence between screen size and input area.

The fifth condition used the tracker in full 3DOF positioning mode. Selection was still done via 2D ray casting, but once selected, objects could be freely moved along all three world axes (without sliding). Collision avoidance was still enabled in this mode. Object movement was directly mapped to tracker position: moving the tracker up caused the object to move upwards in the scene; moving the tracker towards the screen caused the object to move "into" the scene, etc. Speed of object motion in this condition was set to be virtually identical to the other conditions (excluding the large area tracker condition).

4.2.4 Procedure

Participants were first introduced to the experiment and signed consent forms. They were then given a practice trial to familiarize themselves with the task. In addition, they were given verbal feedback throughout the experiment until they were able to remember the task without aid (typically within 2 or 3 trials). The task was the same as in the previous experiment.

Since practically all participants rotated the scene to an overhead view in the first study, we set this as the default viewpoint and disabled scene rotation in this study. Following completion of the experiment, participants were surveyed for subjective preferences.

4.2.5 Design

The study was a 5×6 within-subjects design. The first factor was input technique and the second was trial number. Five input techniques were compared: mouse, "mouse emulation", "large area mouse emulation" (30x22.5 cm mapping), "air-mouse emulation" (as mouse emulation but without support), and 3DOF mode. Note that the "mouse emulation" mode was identical to the supported horizontal device condition from the first study. Similarly, the "air-mouse emulation" mode was identical to the unsupported horizontal device condition from the first study.

Participants performed a total of 30 trials each. In total, it took them approximately 1 hour to complete the experiment.

4.2.6 Results

The dependent variables were again task completion time and accuracy. ANOVA showed a significant difference in task completion time between the five conditions, ($F_{4,295}$=61.19, p<0.0001). Tukey-Kramer post hoc analysis indicated that the mouse condition was significantly faster than all other conditions. All three of the 2D tracker conditions were not significantly different from one another. Finally, the unconstrained 3DOF tracker condition was significantly slower than all others. The mean times for these conditions are visualized in Figure 6.

A significant difference was found in accuracy between the five conditions ($F_{4,290}$=4.65, p<0.005). Tukey-Kramer revealed that the mouse and mouse emulation conditions were significantly more accurate than the 3DOF condition. However, no other conditions were significantly different. Figure 7 summarizes the mean error distances for each condition.

This time, participants clearly preferred support, with an average of 1.4 on a 5 point Likert scale (1 being best). Ranks for the 5 movement techniques were analyzed with a Kruskal Wallis ANOVA and were found to be significantly different (H_4 = 12.52, p<.05), with mean preference scores of 1.6 for the mouse, 3 for "mouse emulation", 3.6 for "air mouse emulation", 3.3 for "large area mouse emulation", and 3.5 for the 3DOF tracker condition. Post-hoc analysis revealed that preference for the mouse technique was significantly higher than all other techniques, with the exception of the "mouse emulation" technique. There was no significant difference in preference between the remaining three techniques.

4.3 Overall Discussion

As discussed above, one concern in the first study was that allowing scene rotation might have confounded the design. Participants might have been moving objects from different screen locations. To further address this, we analyzed two conditions that were present in *both* studies: "mouse emulation" and "air mouse emulation". If the viewpoint rotation had confounded the results, we might be able to see this reflected as significant differences between the identical conditions across experiments. However, comparing all trials for these conditions indicates that neither speed ($F_{3,244}$=1.03, p>.05) nor accuracy

($F_{3,243}$=0.47, ns) were significantly different. Analyzing only corresponding unsupported conditions and supported conditions also fails to show any significant differences.

As the second study had two more trials than the first, the additional learning may have resulted in better performance. To account for this, these analyses were repeated on only the first 4 trials. Again, one-way ANOVA showed no significant difference in speed ($F_{3,204}$=1.38, p>.05) or accuracy ($F_{3,203}$=0.13, ns). Also, neither the "air-mouse emulation" nor the "mouse emulation" conditions showed any significant differences across experiments.

Given that scene rotation time was small compared to the overall times and that we failed to find any significant differences between identical conditions across studies, we hypothesize that scene rotation probably did not confound the first study.

Another issue is that the complexity of the task used in both studies increased the variability, thus making it harder to detect significant differences between conditions. As discussed, we selected the task to improve the external validity of the results – perhaps at the cost of internal validity. However, participants were given a "recommended" ordering of object movements during practice, and almost all adhered to it. Additionally, when they showed signs of confusion as to which object to move next, the experimenter would provide verbal instructions according to the recommended ordering. All of this leads us to believe that our results still address major aspects of our research goals.

4.3.1 Physical Support

The lack of effect for support appears to contradict previous findings [10, 20]. However, one difference is that previous work [10] used a two-dimensional task: direct manipulation of 2D shapes in a plane. Moreover, unlike other previous work [20], the input space in our experiment was disjoint from the display area, which is characteristic of the mouse condition. This is also a feature of the Bat input device, which matches relative movements of the input device to virtual object movement [21]. We attribute the difference in our results to these factors. We hypothesize that a different input strategy that registers the display with the input device (e.g. a stylus/touch-screen) may benefit more from support compared to unregistered approaches.

Another possible explanation for the lack of differences is that the 2D sliding movement technique used here made the 3D movement task equally difficult (or easy) for all input conditions in the first study. Thus, the sliding technique may have had much more influence on the results than any of the investigated factors. This is supported by previous work [19], which reported "three-tiered" results similar to the second study: tracker conditions using the sliding movement technique were better than the 3DOF technique, with both being outperformed by the mouse. However, it is important to realize that a cross-device comparison with *different* input mapping techniques evaluates also the techniques!

The subjective findings from the first study suggested that participants were undecided as to the benefits of support. Comments made by participants ranged from "I didn't like vertical support at all" and "Support felt a bit stubborn" to "Lack of support didn't seem to affect the results" and "Unsupported conditions were uncomfortable". However, users clearly preferred support in the second study, as well as combinations of conditions that more closely resemble a desktop environment. Since these conditions performed best, this is more in line with previous findings about the benefits of support.

4.3.2 Equipment Differences

The extensive familiarity of people with the mouse must be considered. Prior to the using 2D constrained tracker conditions for the first time, participants were warned that although the device felt (physically) like a mouse, it did not behave quite like one: the tracker used absolute positioning, and thus did not require clutching. Participants sometimes tried to clutch to move the cursor more quickly but this had no effect since the device tracked equally well on or off the table. Clutching occurred most often in the large area tracker condition in the second study. This is a potential reason why the large area tracker condition did not perform as well as the mouse emulation, despite the increased relative spatial resolution. However, as the control-display (C-D) ratios for the conditions were the same and input was linear (i.e. no acceleration), one would not expect a difference, see e.g. [12]. Another potential reason is that the differences are due to variations in muscle usage for the larger interaction area, but as the range of motions is not that different, this is also improbable.

The main motivation behind including a large tracking area condition in the second study was a concern about the potential effects of resolution. According to specifications, the IS900 offers 0.75mm resolution, which translates to 200 samples inside a 15cm distance. This was mapped to 800 pixels on the screen. This mismatch in resolution may have degraded performance of the 3D tracker relative to the mouse. In practice, the tracker delivers a bit better precision, so this is a conservative estimate. However, the mouse has a much higher tracking resolution than a 3DOF tracker. Optical mice offer between 400-1600 dpi, which corresponds roughly to 0.05-0.01mm resolution, i.e., between one and two orders of magnitude better than the tracker.

This difference in tracking resolution is arguably the most plausible explanation for the outcome of the second study. The overall familiarity of users with the mouse, the presence/absence of support and differences in how the devices moved are much less probable, but cannot be ruled out. Most likely due to the relative unfamiliarity, the unconstrained 3DOF tracker mode showed the strongest learning effects in the first few trials. An ANOVA was performed to determine after which trial participants no longer improved significantly. The last significant improvement in speed occurred between trials 2 and 3 ($F_{1,18}$=4.41, p<.05). In other words, starting with the 3rd trial there were no observable learning effects and the learning curves effectively flatten off even for the 3DOF mode. Although it is impossible to predict long-term learning effects from only 6 trials, the evidence suggests that it is unlikely that more training would allow the 3DOF mode to match the other conditions without extensive, long-term training.

4.3.3 Muscle Groups

To avoid confounds, we used the same "top-down" grip on the mouse, with the tracker on the wrist in all conditions. Such confounds could arise if, for example, a 3D wand input device was used in the unsupported conditions. This is because different muscle groups would be used to perform motions, since one typically holds a wand-type input device with the hand rotated ~90° relative to how one holds a mouse. This is also supported by previous work [1, 24], which showed that using different muscle groups affects performance in 6DOF docking [24] and Fitt's tasks [1]. Since our experimental task was made up of several of these simple motions, differences between devices would likely be exaggerated. Consequently, we used the same device combination throughout the experiments to ensure that (approximately) the same muscle groups were used in all conditions, and thus provide a more level playing field.

One participant pointed out that they noticed they moved the mouse with their fingers for fine motions. Since the tracker was mounted on the back of the hand, fine motor control motions, such as adjusting the mouse with the fingertips, were unlikely to have been recorded. This may also account for the differences found.

5 Conclusions

We conducted two studies comparing factors affecting the choice of input devices for constrained 3D positioning tasks. The first of these studies compared the effects of matching or mismatching device movement plane to the orientation of the screen, and the presence or absence of a supporting surface. To our surprise, no significant differences were found between these conditions.

The second study compared the mouse to a 3DOF tracker in several conditions. The tracker conditions included a mouse emulation mode with and without support, as in the first study. A larger area mouse emulation mode with support and a 3DOF movement condition were also included. The results show a significant difference in speed between the mouse and all tracker conditions for speed and accuracy. The mouse performed best, followed by the mouse emulation mode. The 3DOF tracker mode performed worst, with the remaining constrained tracker modes in between. These results lead us to conclude that 2D-based movement techniques can be effectively used with 3D devices such as trackers. In our second study, a sliding based movement technique operated with a 3D tracker consistently outperformed a full 3DOF movement technique, even with collision avoidance.

However, the mouse outperformed all tracker conditions. Given the state of current tracking technologies, our results lead us to recommend that for fine-grained manipulation, designers should consider the use of the mouse, tablet, and touch-screen/pen based systems as current 3D trackers simply cannot track as precise.

5.1 Future Work

We are interested in studying other input devices to further assess which properties lend themselves to intuitive 3D manipulation interfaces. In particular, we intend to look at high precision 3D input devices, such as the Phantom. Such a study may help to determine how important tracking precision really is, but one has to account for the different grip and working space. A related avenue for future research is further analysis of the differences between muscle groups used to operate various devices. In particular, if accurate finger tracking in free air could be achieved, would this improve performance to mouse-like levels? We also plan to investigate tablets, as these devices provide high precision and are well suited to the sliding 3D movement technique.

A final area for future research is to examine the effect of scene orientation compared to display and device orientation. While designing the first experiment, we considered including scene orientation (e.g. top-down view vs. side view) as a factor. However, as the experiment was becoming too large, we chose to exclude it. We intend to revisit this in the future.

6 Acknowledgements

Thanks to John Bonnett for use of the lab and equipment, Vicky McArthur for help with the figures and Andriy Pavlovych for help with the video.

References

[1] R. Balakrishnan and I. S. MacKenzie. Performance differences in the fingers, wrist, and forearm in computer input control. *Proceedings of CHI '97*, pp. 303-310, 1997.

[2] E. Bier. Skitters and jacks: interactive 3D positioning tools. *Proceedings of Symposium on Interactive 3D Graphics*, pp.183-196, 1987.

[3] J. Boritz, and K. S. Booth. A study of interactive 3D point location in a computer simulated virtual environment. *Proceedings of VRST '97*, pp. 181-187, 1997.

[4] J. Boritz, and K. S. Booth, A study of interactive 6DOF docking in a computerised virtual environment. *Proceedings of the Virtual Reality Annual International Symposium*, pp. 139-146, 1998.

[5] D. Bowman, D. Johnson, and L. Hodges. Testbed evaluation of virtual environment interaction techniques. *Proceedings of VRST '99*, pp. 26-33, 1999.

[6] J. Chen, M.A. Narayan, and M.A. Perez-Quinones. The use of hand-held devices for search tasks in virtual environments. *IEEE VR 2005 workshop on New Directions in 3DUI*, pp. 15-18, 2005.

[7] D. Brookshire Conner, S. Snibbe, K. Hemdon, D. Robbins, R. Zeleznik, and A. van Dam. Three dimensional widgets. *Proceedings of Symposium on Interactive 3D Graphics*, 1992.

[8] Y. Kitamura, A. Yee, and F. Kishino. A sophisticated manipulation aid in a virtual environment using dynamic constraints among object faces. In *PRESENCE*, 7 (5), pp. 460-477, 1998.

[9] L. Kohli, and M. Whitton. The haptic hand: providing user interface feedback with the non-dominant hand in virtual environments. *Proceedings of Graphics Interface*, pp. 1-8, 2005.

[10] R. W. Lindeman, J. L. Sibert and J. K. Hahn. Hand-held windows: towards effective 2D interaction in immersive virtual environments. *Proceedings of IEEE VR'99*, pp. 205-212, 1999.

[11] D. W. Martin. Doing Psychology Experiments 5th Edition. Wadsworth / Thomson Learning, Belmont CA, 2000.

[12] I. S. MacKenzie and S. Riddersma. Effects of output display and control-display gain on human performance in interactive systems. *Behaviour & Information Technology, 13*, pp. 328-337, 1994.

[13] R. McMahan, D. Gorton, J. Gresock, W. McConnell, D. Bowman. Separating the effects of level of immersion and 3D interaction techniques. *Proceedings of ACM VRST '06*, pp. 108-111, 2006.

[14] J.-Y. Oh, W. Stuerzlinger. Moving objects with 2D input devices in CAD systems and desktop virtual environments. *Proceedings of Graphics Interface*, pp. 195-202, 2005.

[15] I. Poupyrev, N. Tomokazu, and S. Weghorst. Virtual notepad: handwriting in immersive VR. *Proceedings of IEEE VR'98*, pp. 126-132, 1998.

[16] I. Poupyrev, S. Weghorst, M. Billinghurst, T. Ichikawa, Egocentric object manipulation in virtual environments: empirical evaluation of interaction techniques. *Proceedings of Eurographics '98*, pp 41-52, 1998.

[17] G. Smith, T. Salzman and W. Stuerzlinger. 3D scene manipulation with 2D devices and constraints. *Proceedings of Graphics Interface*, pp. 135-142, 2001.

[18] Z. Szalavári, and M. Gervautz. The personal interaction panel, a two-handed interface for augmented reality, *Proceedings of Eurographics '97*, pp. 335-346, 1997.

[19] R. Teather and W. Stuerzlinger. Guidelines for 3D positioning techniques. *Proceedings of Futureplay '07*, 61-68, 2007.

[20] Y. Wang, C. L. MacKenzie. The role of contextual haptic and visual constraints on object manipulation in virtual environments. *Proceedings of CHI 2000*, pp. 532-539, 2000.

[21] C. Ware and D. R. Jessome. Using the bat: a six-dimensional mouse for object placement. *Proceedings of Graphics Interface*, pp. 119-124, 1988.

[22] C. Ware, and K. Lowther. Selection using a one-eyed cursor in a fish tank VR environment. *In ACM Transactions on Computer-Human Interaction*, 4, 4, pp. 309-322, 1997.

[23] D. Wigdor, C. Shen, C. Forlines, and R. Balakrishnan. Effects of display position and control space orientation on user preference and performance. *Proceedings of CHI '06*, pp. 309-318, 2006.

[24] S. Zhai, P. Milgram, and W. Buxton. The influence of muscle groups on performance of multiple degree-of-freedom input. *Proceedings of CHI '96*, pp. 308-315, 1996.