UNCHECKED PREPRINT

Forthcoming in Presence: Teleoperator and Virtual Environments, 2005
Anthony Steed (A.Steed@cs.ucl.ac.uk), Chris Parker
Department of Computer Science, University College London

Evaluating Effectiveness of Interaction Techniques across Immersive Virtual Environment Systems

Abstract

This paper presents the results of a formal experiment to compare different interaction techniques across two types of immersive display: an immersive projection technology (IPT) and a head-mounted display (HMD). Our aim is to investigate the effectiveness of two widely used interaction metaphors, virtual hand and ray casting on these two display technologies. Our motivation is that design and evaluation of interaction techniques for immersive egocentric display systems has been undertaken almost exclusively on HMDs. We argue that basing interaction for IPTs on techniques developed for other types of immersive systems is a flawed approach, as there are some categorical differences between the experience given by an IPT and a HMD. For example, an IPT user has a much wider field of view than a HMD user.

We have chosen two types of interaction tasks to study: simple selection of objects both near to and at some distance from the user, and manipulation of objects involving a change of both position and orientation. As previous studies have found, we find that ray casting is preferable for selection and virtual hand is preferable for manipulation for a HMD. We show that this is also the case for the IPT. More interestingly, whilst we find performance on selection tasks is much better on the IPT, for manipulation tasks there is little difference between the two display technologies.

1 Introduction

Immersive virtual environment (IVE) systems such as head-mounted displays (HMDs) and immersive projection technologies (IPTs) are finding increasing use in scientific and engineering domains (Stone 2002). In fact in recent years, IPT displays such as the CAVE™[1] have become relatively prevalent and are eclipsing HMDs in many laboratories. However relatively little is known about design and engineering trade-offs in selecting a particular type of IVE system, or the interaction techniques that underpin the applications. Two of the field's many "hot open challenges" are choosing the appropriate display for the task and interacting efficiently within the virtual environment (Brooks 1999). This paper attempts to give initial answers to both of these challenges.

One key factor that will determine which immersive technology is chosen is cost. IPTs are relatively expensive compared to HMDs, and both are still relatively expensive compared to desktop displays. There are also numerous health and safety issues with immersive displays (Stanney et al. 1998), though as of yet, there are no direct comparisons between IPT and HMD technologies.

If cost and health and safety issues can be addressed the choice between HMD and IPT depends on the appropriateness of the technology to the application and general issues of usability and utility of the displays (Bowman et al. 2002). In this paper we focus on the low level tasks of selection and manipulation in a virtual environment. This is an area that has received a lot of attention in the field

---

[1] CAVE is a registered trademark of the Board of Trustees of the University of Illinois

(see Section 2.1), but as far as we are aware there are no direct comparisons between HMD and IPT for these tasks.

In this paper we will study two widely used interaction metaphors for selection and manipulation: virtual hand and ray casting. We will look at these interaction metaphors on both a HMD system and an IPT system. We will show the relative effectiveness of the combination of display devices and interaction techniques at selection and manipulation tasks. We will then generate some guidelines about which interaction techniques are to be preferred for combinations of display device and task. In particular we will see that for highly focused manipulation tasks, there is not much advantage to using an IPT. However the IPT provides superior performance at selection tasks. Overall this leads us to suggest that on the majority of applications involving significant interaction and manipulation, an IPT should outperform a HMD.

2 Interaction within Immersive Virtual Environments

The community of virtual environment practitioners still finds it difficult to answer questions about the appropriateness of displays and interaction configurations for specific tasks. Even the choice between a desktop and an IVE system is difficult to make. Comparisons have been made between IVEs and desktop displays (Pausch et al. 1997, Schroeder et al. 2001, Bowman et al. 2002) and they tend to show an advantage to immersion in some applications. This paper will not address the problem of choosing an IVE over a desktop and we will assume that there is a class of applications that demand full immersion. Thus we will only be concerned with variations across different types of IVE system.

IVE systems attempt to *exclude* the real world, so that the virtual environment is perceived as a continuous and consistent environment. In IVE systems the virtual environment display *surrounds* or appears to surround the user. Furthermore, IVE systems also *include* the user in the virtual environment. This means that the system tracks the body of the user and uses this information as an aid to interaction in the system. On a HMD, this tracking information is often given a geometric form, and is rendered as an avatar. Thus as the user looks around the environment and at their own body, the avatar is updated to reflect the actual body movements made.

IVEs afford egocentric displays and body-centric interaction metaphors that enable a wide range of tasks to be performed in a way that mimics similar real world tasks (Mine et al. 1997). Such systems are finding use in many applications where transfer of reactions from the real to the virtual is important for the task at hand. An example, is in training applications, where it is important that the participants understand and accept the presented events in a coherent and believable continuum, and react accordingly (Higgett & Bhullar 1998). It is also widely believed that IVEs are unique in some respects in that they can create a sense of *presence* (Draper et al. 1998, Slater & Steed 2001). It is claimed that this sense of presence is necessary for many applications where the user should feel and react as if they are in the place depicted by the displays.

Fundamental to the design of interaction in such environments is the issue of how the participant moves through the environment and how they select and manipulate objects. Since the user is immersed and tracked, we can use the position of the user's body to determine interaction. However, there are limitations to tracking such as accuracy and range of devices and physical constraints on movement in the real-world, so inevitably some form of interaction metaphor is required. With immersive interaction the issue then becomes how intuitive body-centered interaction is reconciled with the often counter-intuitive interaction metaphors required.

2.1 Selection and Manipulation

A complete review of current immersive interaction techniques is beyond the scope of this paper. Review material can be found in many papers in the area including (Bowman & Hodges 1999, Poupyrev et al. 1997, Mine et al. 1997, Bowman et al. 2001). Bowman et al. 2001 provide a taxonomy of interaction tasks, including selection and manipulation. We will thus only describe a few critical issues in the choice and implementation of selection and manipulation techniques.

Selection within an IVE system is commonly effected by collision between the desired object and the user's virtual hand, or by intersection of the object with a ray projecting in the direction the hand is pointing. The former is more intuitive since it is analogous to the real world, where a user simply reaches out and touches the object. However it does require objects to be within reaching distance. The latter is more flexible but less natural since selection takes place at a distance. More specialized techniques include Go-Go selection (Poupyrev et al. 1996). In Go-Go selection the user’s arm is dynamically extended to reach distant objects. Poupyrev et al. 1998 note that Go-Go is a superset of the virtual hand and whenever the virtual hand is used, Go-Go selection is a natural and flexible extension.

Manipulation is effected by selecting an object to grab it and then moving the hand to specify the new position and rotation. Usually there is a rigid attachment between the hand and the object being manipulated (Mine et al. 1997). Again there are a number of variations on these techniques have been investigated. After selection has been achieved, Go-Go allows the user to manipulate the objects as if their arm had been extended to that object (Poupyrev et al. 1996). HOMER is a similar technique. First an object selected by projecting a ray, and then the object is manipulated as if the arm had been extended to the selected object (Bowman & Hodges 1997). With non arm-extension manipulation techniques, if a close-by object is selected (e.g. by collision detection) then the user can accurately re-position the object in three dimensions and also re-orient the object freely. However the reach of the user limits the distance over which re-positioning can occur. If the object is remotely selected then small motions of the hand cause large movements of the object. This means that moving objects over long distances is easier, but that precise positioning is more difficult. It also means that rotating an object in place may be more difficult, since it can require large or repeated arm movements.

We consider selection and manipulation as two separate atomic interaction techniques. This distinction is made because, although selection can be considered as a sub-case of manipulation, selection will often be done without consideration of subsequent manipulation. When selecting an object, there is no need to consider the direction in which it will be moved, and thus no need to adapt the selection gesture accordingly. For example, there is no need to consider whether the hand can subsequently be moved without strain to the user and without colliding with the physical or virtual environment.

In our selection and manipulation task we will require locomotion to be enabled. Locomotion is commonly effected with a flying metaphor such as "fly in the direction of gaze when a button is pressed”. Any such technique must compete with the natural urge to actually walk. Indeed we frequently observe users in HMDs and IPTs taking a few paces, before remembering (or being reminded by collision with the real world) that to locomote over large distance they must press a button.

2.2 Considerations for IPT Systems

Almost all of the techniques described in the previous section have been designed for or evaluated on HMD systems. There are a few differences between IPT and HMD that deserve discussion.

Since the IPT user can see their own body and can't see a virtual body, they may not see a virtual object attached to the real hand’s position since their own hand may obscure the relevant part of the screen. The virtual hand technique uses collision detection of the geometry of the virtual hand with objects in the scene as the basis for selection and manipulation. Since the virtual hand is located at the tracked position of the real hand, in an IPT implementation, it may be better not to draw the virtual hand since the real hand will usually obscure it.

In a HMD, if there is any significant latency in the system, the user will actually see the virtual hand in a different position to where their proprioception indicates their real hand is. However in an IPT the effect of latency is not perceived directly since usually no virtual hand is rendered and thus the user might believe that selection is effected by touching an object with their real hand. This is subtly wrong, since the real selection volume will lag behind slightly.

The most notable difference between an IPT and a HMD is that the field of view of the former is much larger. However since the most common configuration of an IPT has only three walls and a floor, the virtual environment does not completely surround the user. Thus a joystick is usually provided to rotate the virtual coordinates of the IPT about the user. A two-axis joystick is usually used and the second dimension is mapped to forwards and backwards locomotion. However we have observed users forgetting to turn using the joystick and turning so far that they face a blank wall at which point they remember to use the joystick.

Both the HMD and IPT displays have a limited working volume. However the first a HMD user will usually know about the limited volume is when they walk into an object or get tangled in cables. IPT users do occasionally collide with the walls, but cables are not so much of a problem with a three-sided system since users rarely turn all the way around past the outside wall.

Finally, an IPT is a multi-user system. Usually only one user can be tracked and the view is only correct for that user. This suggests that there are two possibly conflicting primary goals in designing interaction techniques for an IPT: to optimize for the tracked user, or to optimize with consideration of the comprehension of the group of users. In this paper we will focus on the former, and all our experiments will be single user. We have pursued the situation of having a non head-tracked user in another experiment (Steed & Parker 2004).

2.3 Evaluation Strategies

Despite the large number of IVE applications and demonstrations, relatively little attention has been paid to evaluation of the underlying interaction techniques. An overview of strategies for evaluation of virtual environments has been provided by Bowman, Gabbard, & Hix (2002). They discuss the difference between general usability engineering of applications and usability engineering of interaction techniques. Applications are usually optimized to support a specific set of tasks, but interaction is usually optimized to support a generic class of tasks that support a wide range of applications. There is a tension between designing for a specific application and generalizing to other applications.

To counter this, Lampton et al. (1994) and Bowman et al. (2001) have proposed test-beds for interaction tasks that contain a battery of standard interaction tasks. We have chosen not to take either of these test-beds, but rather to focus on a relatively small number of task types with a relatively large number of repeated trials within each task (see Section 3.3). The main reason for this is that test-beds contain many tasks, and our main aim in this study is to highlight differences between IVE displays systems and between virtual hand and ray casting. We are not studying the effectiveness of virtual hand or ray casting, nor are we studying a specific application. A complete battery of generic tasks would have been unnecessarily detailed at this stage.

In this paper we are concerned with selection and manipulation strategies, but note that locomotion has been studied for HMD versus IPT (Bowman, Datey et al. 2002). In that study, it was found that locomotion is more natural in a HMD than an IPT. This seems to be because of their IPT's missing a back wall. This caused the participants to look partially out of the IPT when turning abruptly, which disrupted the user and forced them to switch to using a joystick to turn the view.

3 Experimental Design

3.1 Equipment

The IPT used in these trials was a Trimension ReaCTor, with three back-projected 2.8x2.2m walls, and a front projected 2.8x2.8m floor. The HMD condition used a Virtual Research V8 helmet with 60 degree diagonal field of view in stereo. Both of these display systems are typical of the types of devices used in the IVE community. Figure 1 shows a HMD user inside the IPT space. Figure 2 shows an IPT user performing a typical task during the experiment.

<Figure 1 HMD user standing inside the IPT space>

<Figure 2 IPT user undertaking a selection and manipulation trial>

Visual output was produced by a SGI Onyx2 with eight 300MHz R12000 processors, 8GB ram and four InfiniteReality2 pipes. For the IPT display, all four pipes were used and stereo was achieved using shutter glasses. A full stereo image was thus generated for the user at 45Hz with the projectors running at 90Hz at a resolution of 1024x768. For the HMD, one pipe was used with two viewports. A full stereo image was generated at 60Hz at 640x480 resolution per eye. Thus the frame-rate was higher for the HMD, but the screen resolution was lower.

The tracking system used was an Intersense IS-900. Users either had the tracking unit attached to the top of the stereo glasses or rigidly fixed to the HMD. HMD users stood inside the ReaCTor so that the working volumes would be the same in both conditions. Both HMD users and IPT users held a wand controller with four buttons and an isotonic joystick. Only one button was used in the experiment. The joystick was used in some tasks. The Intersense configuration for the ReaCTor has all of the transmitter bars on the ceiling, which means that if the hand-held device is not visible from the ceiling, tracking may be lost.

The experimental system software was implemented in the VRJuggler framework (Bierbaum et al. 2001). We used VRJuggler Version 1.1DR2 and customized it to interface to the trackd software provided by Intersense as the device driver for the tracking system. We used the OpenGL Performer library for rendering.

3.2 Interaction Techniques

Despite a large amount of research in selection and manipulation techniques (see Section 2.1) in practice the virtual hand (VH) and ray casting (RC) techniques are still used far more commonly than any other technique. These are thus the most natural techniques for us to study when comparing HMD and IPT displays.

Ray casting selection was implemented using the OpenGL Performer ray casting functionality. The ray was centered on the hand and cast in the direction of pointing. A thin cylinder was drawn around the ray, but no other feedback was given. Performer does not natively support full object-object intersection so the virtual hand selection was implemented by using a cluster of twelve short rays centered on the hand. In both ray casting selection and virtual hand selection, the user had to press a button on the wand controller in order to indicate their selection.

In the manipulation tasks, the user had to indicate the object in the same was as they did in ray casting or virtual hand selection. The object would be grabbed when the button was pressed and released when the button was pressed a second time. During the grab phase the object was rigidly attached to the hand. A single manipulation trial was expected to take several iterations of grab and drop.

3.3 Tasks

Previous work has indicated the effect of target distance and target size on the performance on selection and manipulation (Poupyrev et al. 1998). We varied target distance but decided not to vary target size because of the relative difficulty of manipulating very small or very large objects.

For the selection tasks we disabled locomotion. Because we wanted to evaluate the virtual hand technique for selection, on one task we restricted object placement to inside the IPT space. This is referred to as the near space selection task. For ray casting only we investigated selection performance over a medium distance, that is, further than 3m. Thus there are two selection tasks: Near Space Selection (NSS) task and Medium Space Selection (MSS) task.

As has been noted by several authors including Mine et al. (1997), ray casting is not ideal for manipulation since the object is some distance from the hand (see Section 3.2). However in pilot experiments we found that it was usable for medium space manipulation and of course in near space manipulation it becomes very similar in use to the virtual hand. There is just one manipulation task (SM). Locomotion was enabled for this task

Poupyrev et al. (1997) provide a VR Manipulation Assessment Testbed (VRMAT) for generating task conditions dynamically as the user position changed. In contrast we chose to fix the sequence of trials and we thus generated a suitable sequence covering a suitable set of trials for the task condition just once. This choice was made because we wished to study variance in actual trial completion times. We will see in Sections 4.1 & 4.3, that in the IPT condition, selection times are highly uniform across individual trials, but in the HMD condition they show high variance.

All of the tasks took place inside a virtual kitchen 8m x 6m x 3m high. In the following descriptions of locations of objects, we will use a world coordinate system. World coordinates are aligned with IPT coordinates prior to any locomotion. Note that for the NSS task, where locomotion is disabled, world coordinates and IPT coordinates remain the same throughout. Before any locomotion the center of the floor of the IPT is at (0, 0, 0). As the user faces the central wall, X increases to the right, Y increase upwards, and Z increases out of the IPT. The virtual room contains some static virtual objects to aid orientation.

The tasks comprised the following trials. Note that each subject sees the same sequence of trials within a task:

- NSS: select 39 objects.
- MSS: select 39 objects.

- SM: select and manipulate 9 objects from their original position to a target position.

A single selection trial in the NSS task or MSS task involved the target object appearing, the user's selecting it causing it to disappear, and a pause of one second before the next trial started. The target object was a 0.25m long teapot.

In the NSS task, each object position in the sequence was chosen at random from a volume P_{MIN}(-1.0,0.6,-1.0) P_{MAX}(1.0,1.6,0.5), that is, a box 2m x 1.5m by 1m high centered just in front of the user and with all objects easily reachable without locomotion.

In the MSS task, the each object position in the sequence was chosen at random from a volume P_{MIN}(-3.0,0.6,-3.5) P_{MAX}(2.5,1.6,4.0) , that is, a box 6.5m x 7.5m by 1m high.

A single manipulation trial in the SM task involved the object and target appearing, the user's making several manipulations until the target is reached, the object and target disappearing and a pause of 1s before the next trial started. The object was a teapot with a stone texture. A purple teapot indicated the target position. The target was deemed to have been reached if the object was within 20 degrees along the shortest angle-axis decomposition and the object position was within 0.02 m of the target position. Each starting position and target position in the sequence were both chosen at random from a volume defined by P_{MIN}(-3.0,0.6,-3.5) P_{MAX}(2.5, 1.6, 4.0) , that is a box 6.5m x 7.5m by 1m high. This ensured that all the objects were in the virtual room and were reachable. Objects were constrained so they could not be moved out of the room. See Figure 3 for a screen shot of an object and target in one of the selection and manipulation trials.

<Figure 3 A screen shot of a selection and manipulation trial using ray casting. The stone texture teapot must be aligned with the purple teapot.>

3.4 Method

Forty subjects, thirty-two male, eight female, were recruited from the staff and students of University College London. Each was paid £4 (about $7) for taking part. Each subject used both the virtual hand and ray casting techniques, with half of subjects doing one technique first and half the other. Each subject only experienced one immersive condition (HMD or IPT). The tasks were done in the following order for subjects who experienced virtual hand first:

VH-NSS, VH-SM, RC-NSS, RC-MSS, RC-SM

And in the following order for subjects experienced ray casting first:

RC-NSS, RC-MSS, RC-SM, VH-NSS, VH-SM

Note that the medium space selection task is not possible with the virtual hand since locomotion with the joystick was disabled.

Subjects were trained twice, once with the virtual hand technique immediately before their sequence of virtual hand tasks and once with the ray casting technique immediately before their sequence of ray casting tasks. Each training session involved a short locomotion task, three selection tasks and two manipulation tasks. Before each task they were given a brief overview of the trials they were expected to perform. During this overview they were told whether or not they would be able to locomote and whether they would be selecting or manipulating objects. No subject had problems understanding the instructions.

When the subject stood in the IPT or put on the HMD within the IPT, the virtual lighting would initially be dimmed. Participants were instructed that the trial would start when they indicated they were ready by pressing a button, at which point the lights were raised and the first trial began. After completing all the trials the virtual lights dimmed.

4 Results

For each task the independent variables were display condition and interaction metaphor condition. The dependent variable was the trial completion time. We will present the results for each task in order, and for each task compare across display devices and interaction technique. As mentioned in Section 3.1, occasionally the tracking system failed. When this happened, it could take 20 seconds or more to regain tracking. However these tracking failures were easily characterized from log files and we removed trials where tracking loss occurred.

4.1 Near Space Selection (NSS)

A summary of the time taken to perform the near space selection (NSS) trials can be found in Table 1. There is a very distinct difference between overall HMD and IPT performance, with selections taking longer on the HMD for both interaction techniques.

<Table 1 Summary of experimental results for all three tasks. Mean and standard deviation of task completion time for the three tasks, near space selection (NSS), selection manipulation (SM) and medium space selection (MSS) across display type and interaction metaphor>

Looking at the performance on HMD, we find a significant difference between ray casting and virtual hand ($p < 0.021$). Ray casting was faster than virtual hand with an average of 6.85 seconds versus 7.58. For the IPT, we find a more strongly significant difference that favored ray casting ($p < 0.0001$). Ray casting took an average of 2.23 seconds rather than 2.73 seconds for the virtual hand.

If we look at performance across virtual hand and ray casting conditions, we find that both are significantly faster on the IPT than the HMD, in both cases with $p < 0.0001$. Figures 4 and 5 give the average performance across the individual trials comprising the NSS task.

<Figure 4 Comparing average trial completion time for the near space selection (NSS) with ray casting across display condition>

<Figure 5 Comparing average trial completion time for the near space selection (NSS) with ray casting across display condition>

We note the much larger variance of completion times for the HMD. Indeed the variance for the IPT case is very low, and fairly constant across the individual trials.

4.2 Selection Manipulation (SM)

Table 1 contains a summary of the time taken to perform the selection manipulation (SM) trials. As can be seen, the difference between overall HMD and IPT performance is not large. There is no significant difference between the virtual hand performance on HMD and IPT ($p < 0.076$). In this case the mean is 19.24 seconds for the IPT and 21.08 for the HMD. A larger study may find an increased significance, but the actual effect is not large, being a 2 second or 10% difference. There is a significant difference for the ray casting performance across display conditions, with performance being better on the IPT ($p < 0.017$). The difference is between a mean of 26.30 seconds for the IPT and a mean of 30.29 seconds for the HMD.

Figures 6 and 7 illustrate this point further by examining individual trial completion times. We can see that unlike Figures 4 and 5, the distinction between the two displays is not so great. The tendency is for the IPT to be faster and have a lower variance, but we cannot see the regularity in IPT performance and irregularity in HMD performance that we had in the NSS task.

<Figure 6 Comparing average trial completion time across display condition for selection manipulation (SM) task ray casting>

<Figure 7 Comparing average trial completion time across display condition for the selection manipulation (SM) task with ray casting>

If we instead focus on the displays themselves and compare across interaction technique, we see the distinct difference between performance with ray casting and virtual hand. See Figures 8 and 9.

<Figure 8 Comparing average trial completion time across interaction metaphor for the selection manipulation (SM) task on IPT>

<Figure 9 Comparing average trial completion time across interaction metaphor for the selection manipulation (SM) task on HMD>

On both HMD and IPT, the virtual hand is significantly faster than ray casting (both $p < 0.0001$). The absolute difference between the means is quite large here. For the IPT the virtual hand has a mean of 19.24 seconds versus 26.30 seconds for ray casting. For HMD the virtual hand has a mean of 21.08 seconds versus 30.29 seconds for ray casting.

4.3 Medium Space Selection (MSS)

Table 1 contains a summary of the time taken to perform the medium space selection (MSS) trials. Although we have only one interaction metaphor here, we can see a similar situation to the NSS trials, with ray casting performance being much superior on the IPT. On average selections take 2.40 seconds on the IPT and 3.89 seconds on the HMD.

Figure 10 shows the individual trial completion times. We see a similar effect to Figures 4 and 5, with both a much lower trial completion time for the IPT, and a much more regular variation across trials.

<Figure 10 Comparing average trial completion time across display condition for the medium space selection (MSS) task with ray casting>

5 Discussion

5.1 Comparing the Display Technologies

Selection Tasks

Perhaps the most remarkable difference between the two display technologies is the difference in the characters of their variance on selection tasks. Figures 4, 5 and 10 show that the variance in trial completion time for the IPT is much lower for every trial, and that the variance is much more regular across trials. For comparison, on the HMD we saw high variance and high variation in variance across the individual trails. This held for both ray casting and virtual hand in NSS task, and for ray casting in MSS task. Note that the object placement for each trial is the same for each user, but the sequence was initially generated by a random process.

The low variation in mean and variance across trials for the IPT situation suggests that for an IPT, selection is almost a constant time activity. In both the NSS and MSS tasks objects could appear across and beyond the whole field of view. Despite this, we rarely saw a selection taking more than 5 seconds. In contrast on the HMD, selection could take up to 30 seconds if the object was difficult to locate.

We suggest that this difference might be explained by the field of view. Previous studies have compared field of view effects on interaction tasks. For example (Alfano & Michel 1990) have noted degradation in performance on real world locomotion, search and manipulation tasks when subjects were restricted to a 60 degree field of view. Arthur (2000) reports on a series of experiments on wide field of view HMDs. He found that performance on searching and navigation tasks was significantly impaired on a HMD with 112 degree field compared with a HMD with 176 degree field of view. He also found performance was further impaired when the field of view was restricted to 48 degrees.

The effect of field of view is supported by comparing between the mean completion time of MSS and NSS tasks on the HMD. From Table 1 we note that in the MSS task mean trial completion time was 3.89 seconds, compared to 6.86 in the NSS task. The MSS task is characterized by the objects being in a larger space. Thus, although the objects are further away, they are less likely to be in the extremes of the field of view. With the NSS task, the objects are closer in and often appear close to the body and out of the field of view.

Overall then, we can expect lower performance with a HMD on a variety of applications involving selection and manipulation. Alfano & Michel's experiment was with restricted vision, and not HMDs, though interestingly as one of their conditions they used 60 degrees, a field of view almost identical to that in Virtual Research V8 helmet used in this study. Of course in the V8 helmet we also have a relatively low display resolution. Although we did not vary display resolution, we note that we observed no cases of users failing to recognize objects because of the limited display resolution. HMD users spent a lot more of their time performing visual scans of their environment to locate the target object, and this visual scanning may dominate the time in the selection tasks. We have not determined exactly what time was spent scanning, but users were observed spending a lot of time scanning for an object that was in a position awkward to see on the HMD, such as by their knees. Scanning behavior was not often observed for IPT users. Overall we might suggest that much wider variance in selection performance on the HMD in Sections 4.1 & 4.3 was due to this scanning.

Selection and Manipulation Task

In contrast to selection only, the SM task was performed at roughly the same performance on HMD and IPT. We hypothesize that this is because once the object has been selected and brought into rough alignment with the target; the task is basically the same on both displays since both target and object are within a small field of view. Thus any difference might be explained by the initial time taken to identify the object and target. Although Section 4.3 reported no significant difference between performances, on average the IPT performed 2.9 seconds better, and the variance was lower, following the trend in NSS performance. This leads us to speculate that the performance of the close to target manipulation of the objects was actually very similar across display.

Furthermore if identifying the target and object in the SM task involved two distinct search tasks the difference between HMD and IPT might be explained by the relative difficulty of visual searching with a HMD. Above we speculated that the primary difference in NSS and MSS tasks was the visual search time. In the MSS task, the selection time with ray casting was on average 1.5 seconds faster on the IPT than on the HMD (see Figure 13). In the SM task the selection and manipulation with ray casting was on average 4.0 seconds faster on the IPT. Thus a significant proportion of the difference could again be explained by visual scanning. These speculations would need to be confirmed by further studies where the object and target are both in the field of view when the trial starts. But we suggest that after the object and target have been located, the actual performance of the manipulation actions might be very similar for HMD and IPT displays.

5.2 Comparing the Interaction Techniques

Near Space Selection Task

On the NSS task, for both HMD and IPT, ray casting is superior. This is just significant on the HMD and strongly significant on the IPT. This is as expected because with the virtual hand technique an object may require a short physical motion to reach the object, whereas with ray casting there is no need to move. One observation to make here is that with the virtual hand technique in the IPT, there can occasionally be confusions due to the IPT not being able to display anything between the eyes and the physical hand. Thus when moving the hand over the object, there is a small visual discontinuity where the hand appears to be both in front of the object due to occlusion cues and behind, due to stereo cues. We do not expect that this had a significant impact in this study, but it is worth further attention. In particular it suggests that for applications that involve a lot of working with objects close to the hands, a HMD might be preferred.

Selection and Manipulation Task

On the SM task for both display devices, virtual hand performed better than ray casting. The observation here is that users would often make more pick and drop actions with ray casting in order to move the object close to its target. They would also occasionally try to orient the object at a distance, an action that is intrinsically difficult to perform. These results are with naïve users, and we note informally, that in our experience expert users sometimes prefer ray casting. Expert users are usually good at complex decompositions of tasks into a series of manipulations at a distance and manipulations close to hand. Although we don't note any improvement in performance over the trials, we suggest that over longer periods, people might learn to tackle manipulations in a more flexible way. Again this deserves further exploration.

5.3 Guidelines for Interaction Metaphors and Display Devices

From our discussion in Sections 5.1 & 5.2 we can propose a few guidelines to aid developers choose between a HMD and an IPT and between ray casting and virtual hand techniques.

The first guideline concerns the choice of display device: "*For VE applications involving repeated selection of objects in near and medium space, choose an IPT over a HMD to provide enhanced performance*". From Section 5.1 and 5.2 we cannot generalize to a guideline for manipulation, since the performance of both HMD and IPT was quite similar.

Bowman et al. (2001) propose a guideline for choosing a selection and manipulation metaphor: *"Use ray-casting (two-DOF) techniques if speed of remote selection is a requirement*". This contrasts with a guideline from Poupyrev et al. (1998) *"For VE applications that primarily require selection and manipulation of objects, chose the virtual hand technique over the ray casting technique".*

Our study helps clarify this a little and we propose: *"For applications that primarily require selection without manipulation, chose the ray casting technique over the virtual hand technique"* and we support Poupyrev's suggested guideline though add a qualifier: *"For VE applications that primarily require selection and manipulation of objects and little or no selection without manipulation, chose the virtual hand technique over the ray casting technique"*

Of course applications rarely involve just selection only or just selection and manipulation only. What we can suggest by looking at the average completion time in Table 1 is that since the advantage of ray casting over virtual hand in NSS task is of the order of 1 second and the advantage of virtual hand over ray casting in the SM task is in the order of 10 seconds, we tentatively propose: *"For VE applications that involve mixture of selection and selection and manipulation tasks, unless the number of occurrences of selection actions dominates the selection and manipulation actions by a factor of 10 to 1, chose virtual hand over ray casting".*

5.4 Future Work

In this paper we have highlighted several aspects of interaction that would be interesting to study in further trials. For example, we noted some problems with visual discrepancies when trying to use virtual hand technique to select objects nearby. One of the most significant areas for further study we raised, was that of expert users versus naïve users. This is a topic for which only anecdotal evidence is available. From our own experience we noted that experts often prefer ray casting over virtual hand. This may be because at very short distances it works very much like the virtual hand, but when the object and target are appropriately configured, a large movement can be made efficiently.

Several other areas can be suggested for further study. For example, one distinctive of an IPT system is that can be used by multiple users simultaneously. In another paper, we note that we might prefer ray-casting over virtual hand technique if the IPT is being used by multiple users, especially if the user doing the selection is not head-tracked (Steed & Parker 2004). Another area for study is variation in performance of selection and manipulations for variations of a specific technology. For IPTs, the obvious variation is in the number of walls. However, it has been noted in longitudinal trials of IPTs, that minor differences in the hand controller and brightness of the walls might cause asymmetries in interaction (Heldal et a. 2004). To expand on the first of these: some IPT controllers have joysticks to rotate the IPT; some do not have a joystick and rely on hand gestures to rotate the IPT; some IPTs do not need a rotational control because they have six sides so the user can always face the direction they wish to go.

Finally, we note that we have only considered selection and manipulation of objects that are unconstrained. The results may be different for constrained objects such as menus and icon bars that are commonly used for effecting application control inside IPT applications.

6 Conclusions

This study has compared performance of selection and selection and manipulation tasks across two different display types: an IPT and a HMD, and across two interaction metaphors: virtual hand and ray casting.

Our main conclusions are that for naïve users:

- The performance, as measured by completion time, of selection only tasks was significantly higher on an IPT.
- On an IPT, the performance of selection, as measured by completion time, showed remarkable uniformity across different positions of the target objects.
- The performance, as measured by completion time, of tasks involving both selection and manipulation was only marginally higher on an IPT.
- For selection only tasks, ray casting has superior performance, as measured by completion time, to virtual hand.
- For selection and manipulation tasks, virtual hand has superior performance, as measured by completion time, to ray casting.

Taking these two last points, and the fact that the benefit of virtual hand for selection and manipulation probably outweighs the benefit of ray casting on selection only tasks, we suggested that virtual hand should be preferred for most tasks unless selection only tasks dominate.

We cannot fully answer the question of whether to choose an IPT over a HMD. Indeed there is now conflicting advice since Bowman et al. 2002 conclude with the following guideline "*For VE applications involving navigation through enclosed spaces and frequent turning, choose a HMD …*". Since most applications involve navigation, selection and selection and manipulation, there is now a dilemma. It would seem that for the moment there is no replacement for studying the application, finding the relative importance of each sub-task and implementing a test-bed comparison between the two display conditions. We will note though, that in Bowman et al.'s study the main limitation for navigation on an IPT was the lack of a back wall, and the subsequent disturbances this caused on large rotations. This suggests that although the field of view of a three-walled IPT is good, the lack of a complete surround does detract from the experience.

Of course, if six-walled IPTs or some other completely surrounding technology become prevalent, this issue may go away. In the mean time, Razzaque et al. (2002) have recently proposed a redirection technique that reduces the number of instances of a user's seeing the missing back wall, by imperceptibly rotating the world around them. So far this has only been tested on navigation tasks, and it will be interesting to see if it can be integrated with selection and manipulation. If so, this may lead to the reversal of Bowman's guideline, and then the advantages of an IPT will be clearer.

Acknowledgments

Thanks to David Swapp for help on UCL's ReaCTor system. All the code, data and documentation necessary to run the experiments is available on the web at

http://www.cs.ucl.ac.uk/research/vr/Projects/Interact/

References

Alfano, P.l. & Michel, G.F. (1990) Restricting the field of view: perceptual and performance effects. Perceptual and Motor Skills, 70(1):35-45.

Arthur, K.W. (2000) Effects of Field of View on Performance with Head-Mounted Displays, *PhD Dissertation, UNC Department of Computer Science*, April 2000

Bierbaum, A., Just, C., Hartling, P., Meinert, K., Baker, A. &Cruz-Neira, C. (2002) VR Juggler: A Virtual Platform for Virtual Reality Application Development, Proceedings of IEEE Virtual Reality, Yokohama, Japan, March 2001.

Bowman, D. & Hodges, L. (1997) An Evaluation of Techniques for Grabbing and Manipulating Remote Objects in Immersive Virtual Environments. *Proceedings of the 1997 Symposium on Interactive 3D Graphics*, 1997.

Bowman, D. & Hodges, L. (1999) Formalizing the Design, Evaluation, and Application of Interaction Techniques for Immersive Virtual Environments. The Journal of Visual Languages and Computing, 10(1), February 1999, pp. 37-53.

Bowman, D., Johnson, D. & Hodges, L. (2001) Testbed Evaluation of Virtual Environment Interaction Techniques. *Presence: Teleoperators and Virtual Environments*, 10(1), pp. 75-95.

Bowman, D., Gabbard, J. & Hix, D (2002) A Survey of Usability Evaluation in Virtual Environments: Classification and Comparison of Methods. *Presence: Teleoperators and Virtual Environments*, 11(4), pp. 404-424

Bowman, D., Datey, A., Ryu, Y., Farooq, U. & Vasnaik, O. (2002) Empirical Comparison of Human Behavior and Performance with Different Display Devices for Virtual Environments. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, pp. 2134-2138.

Brooks,Jr., F.P. (1999) What's Real About Virtual Reality? *IEEE Computer Graphics and Applications*,19, 6:16-27.

Draper, J. V., Kaber, D. B. & Usher, J. M. (1998). Telepresence. *Human Factors,* 40(3): 354-75.

Heldal, I., Steed, A., Spante, M., Schroeder, R., Bengtsson, S., Partanan, M. (2004) Successes and Failures in Co-Present Situations, forthcoming in *Presence: Teleoperators and Virtual Environments*, 2004, MIT Press.

Higgett, N. & Bhullar, S. (1998) An Investigation into the Application of a Virtual Environment for Fire Evacuation Mission Rehearsal Training. *In Eurographics 16th Annual Conference*, Leeds, pp. 87-96. 1998.

Mine, M., Brooks Jr., F. P. & Sequin, C. (1997). Moving Objects in Space: Exploiting Proprioception in Virtual-Environment Interaction. *Proceedings of SIGGRAPH 97*, Los Angeles, CA.

Pausch, R., Proffitt, D. & Williams, G. (1997) Quantifying immersion in virtual reality. *In Proceedings of the 24th annual conference on Computer graphics and interactive techniques*, pp. 13–18. ACM Press/Addison-Wesley Publishing Co., 1997.

Poupyrev, I., Billinghurst, M., Weghorst, S. & Ichikawa, T. (1996) Go-Go interaction technique: Non-linear mapping for direct manipulation in VR. *In Proceedings of UIST'96*. ACM. pp. 79-80.

Poupyrev, I., Weghorst, S., Billinghurst, M. & Ichikawa, T. (1997) A Framework and Testbed for Studying Manipulation Techniques for Immersive VR. *In Proceedings ACM Virtual Reality Systems and Technology Conference (ACM VRST'97), Lausanne, Switzerland.*

Poupyrev, I., Weghorst, S., Billinghurst, M. &Ichikawa, T. (1998) Egocentric object manipulation in virtual environments: empirical evaluation of interaction techniques. *Computer Graphics Forum, EUROGRAPHICS'98 Iissue*, 17(3), pp. 41-52.

Razzaque, S., Swapp, D., Slater, M., Whitton, M. & Steed A. (2002) Redirected Walking in Place, *Eighth Eurographics Workshop on Virtual Environments, Barcelona,* May 2002.

Schroeder, R., Steed, A., Axelsson, A-S, Heldal, I, Abelin, A., Widestrom, J., Nilsson, A., Slater, M. (2001) Collaborating in Networked Immersive Spaces: As Good as Being There Together?, *Computers & Graphics, Special Issue on Mixed Realities - Beyond Conventions*, 25(5), October 2001.

Slater, M. & Steed, A. (2000) A Virtual Presence Counter, *Presence: Teleoperators and Virtual Environments*, 9(5), October 2000, MIT Press, ISSN 1054-7460.

Stanney, K.M., Mourant, R. & Kennedy, R.S. (1998). Human factors issues in virtual environments: A review of the literature. *Presence: Teleoperators and Virtual Environments*, 7(4), 327-351.

Steed, A. & Parker, C. (2004). 3D Selection Strategies for Head Tracked and Non-Head Tracked Operation of Spatially Immersive Displays. *Immersive Projection Technology (IPT2004), Ames, Iowa*, 13-14 May 2004.

Stone, R.J. (2002) Applications of Virtual Environments: An Overview, In *Handbook of Virtual Environments: Design, Implementation and Applications,* Stanney, K.M (editor), Lawrence Erlbaum Associates, pp. 827-856.

Usoh M, Arthur K, Whitton M, Bastos R, Steed A, Slater M, Brooks F (1999) Walking > Walking-in-Place > Flying, in Virtual Environments, *Proceedings of SIGGRAPH 99 (Los Angeles, California, August 8-13). In Computer Graphics Proceedings, Annual Conference Series, ACM SIGGRAPH*, pp. 359-364.

Figure 1 HMD user standing inside the IPT space

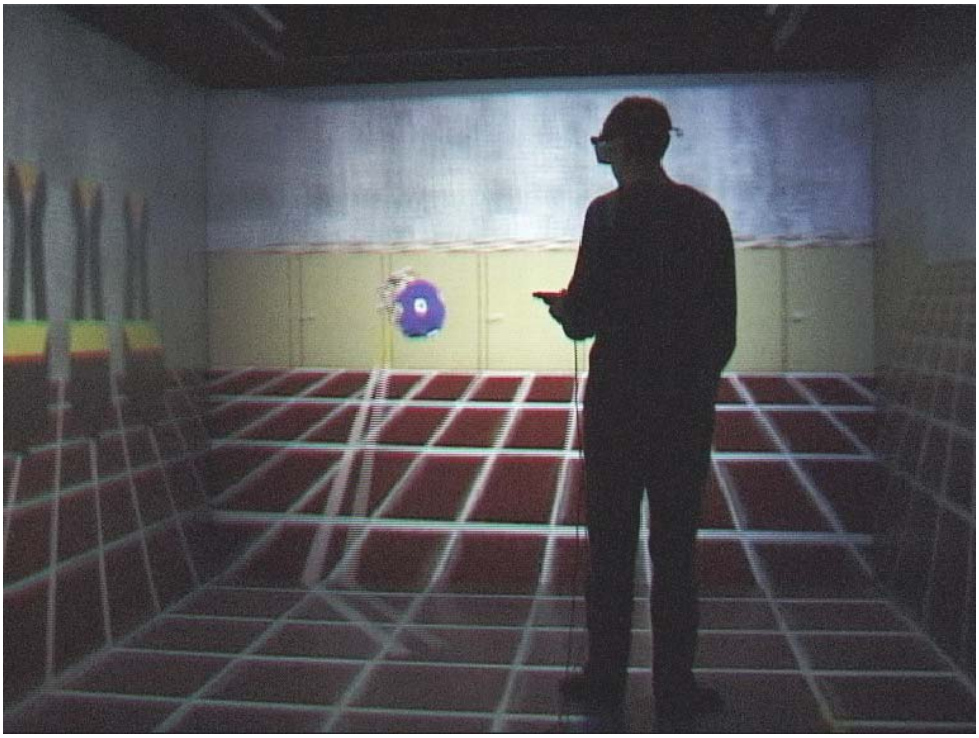

Figure 2 IPT user undertaking a selection and manipulation trial

Figure 3 A screen shot of a selection and manipulation trial using ray casting. The stone texture teapot must be aligned with the purple teapot.

Display	Interaction Technique	Tasks		
		NSS	SM	MSS
HMD	Virtual Hand	7.57 ± 5.76	21.08 ± 10.15	
	Ray Casting	6.86 ± 6.48	30.29 ± 17.39	3.89 ± 2.59
IPT	Virtual Hand	2.73 ± 0.93	19.24 ± 8.97	
	Ray Casting	2.23 ± 0.90	26.30 ± 13.65	2.40 ± 0.63

Table 1 Summary of experimental results for all three tasks. Mean and standard deviation of task completion time for the three tasks, near space selection (NSS), selection manipulation (SM) and medium space selection (MSS) across display type and interaction metaphor

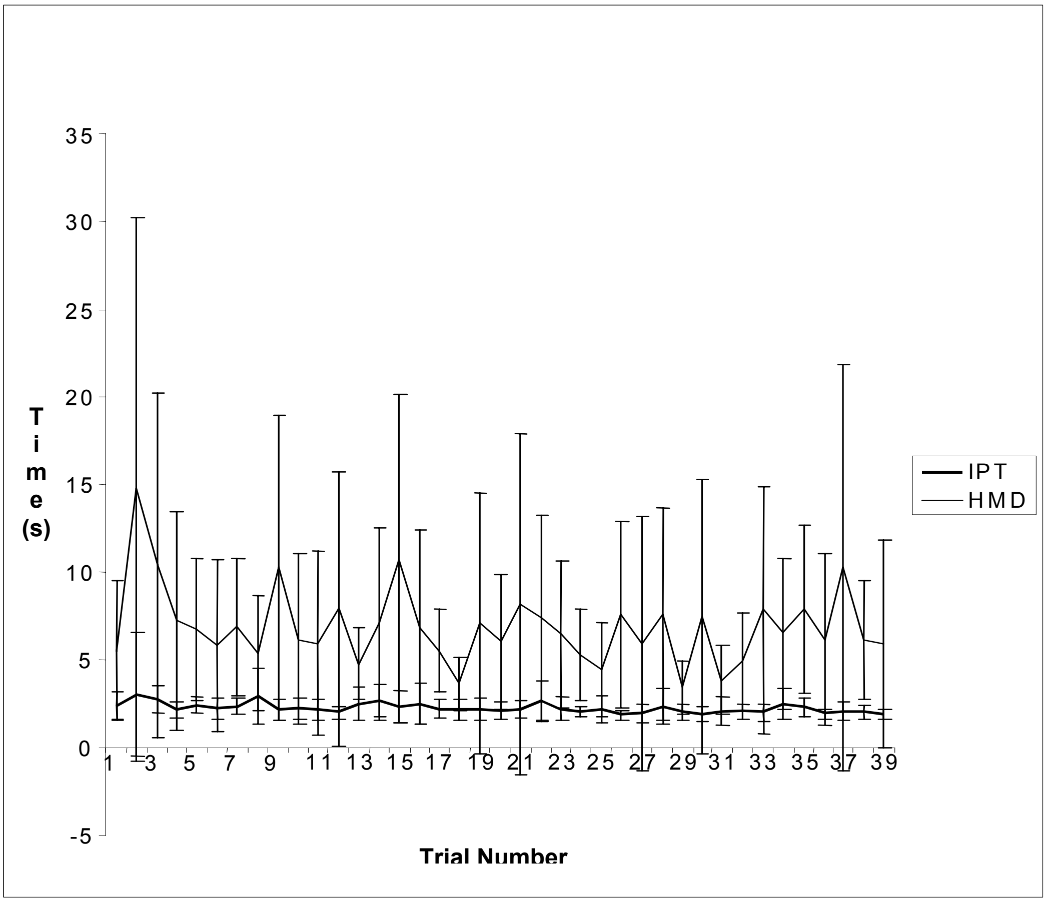

Figure 4 Comparing trial completion time across display condition for the near space selection (NSS) task with ray casting

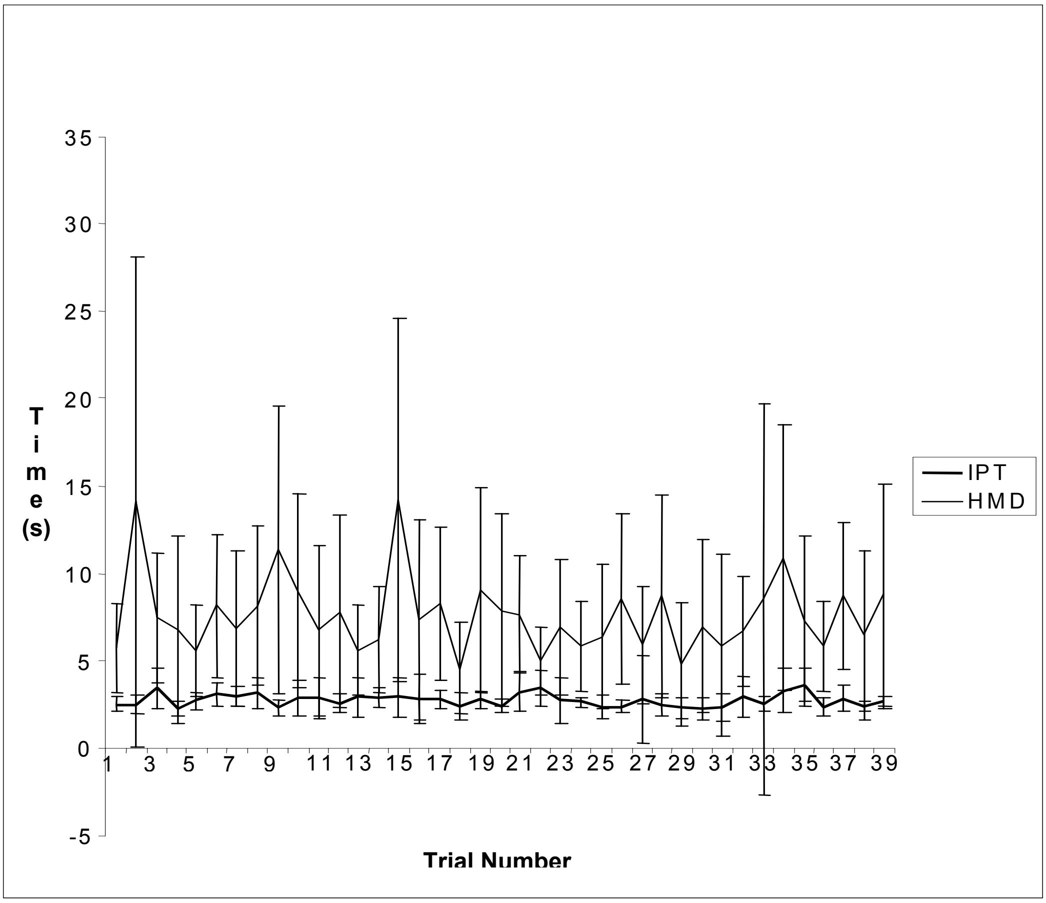

Figure 5 Comparing average trial completion time across display condition for the near space selection (NSS) task with ray casting

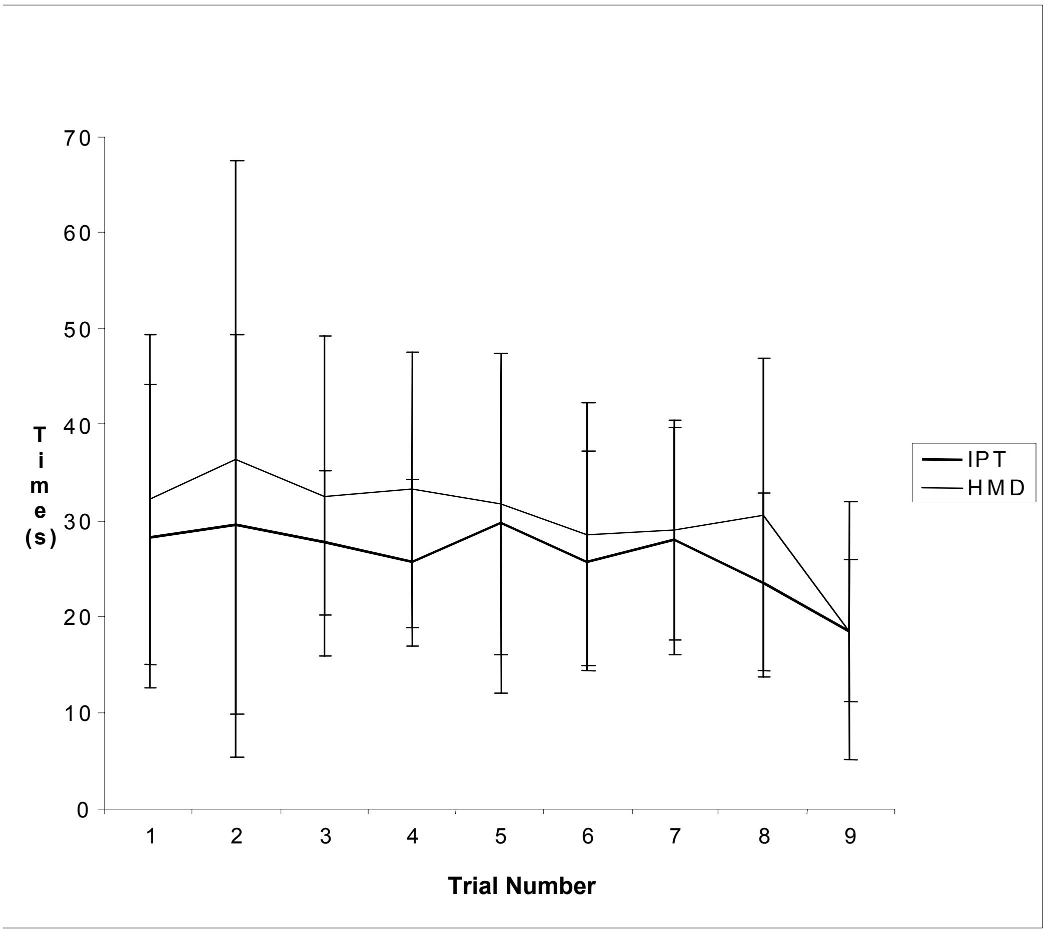

Figure 6 Comparing average trial completion time across display condition for selection manipulation (SM) task ray casting

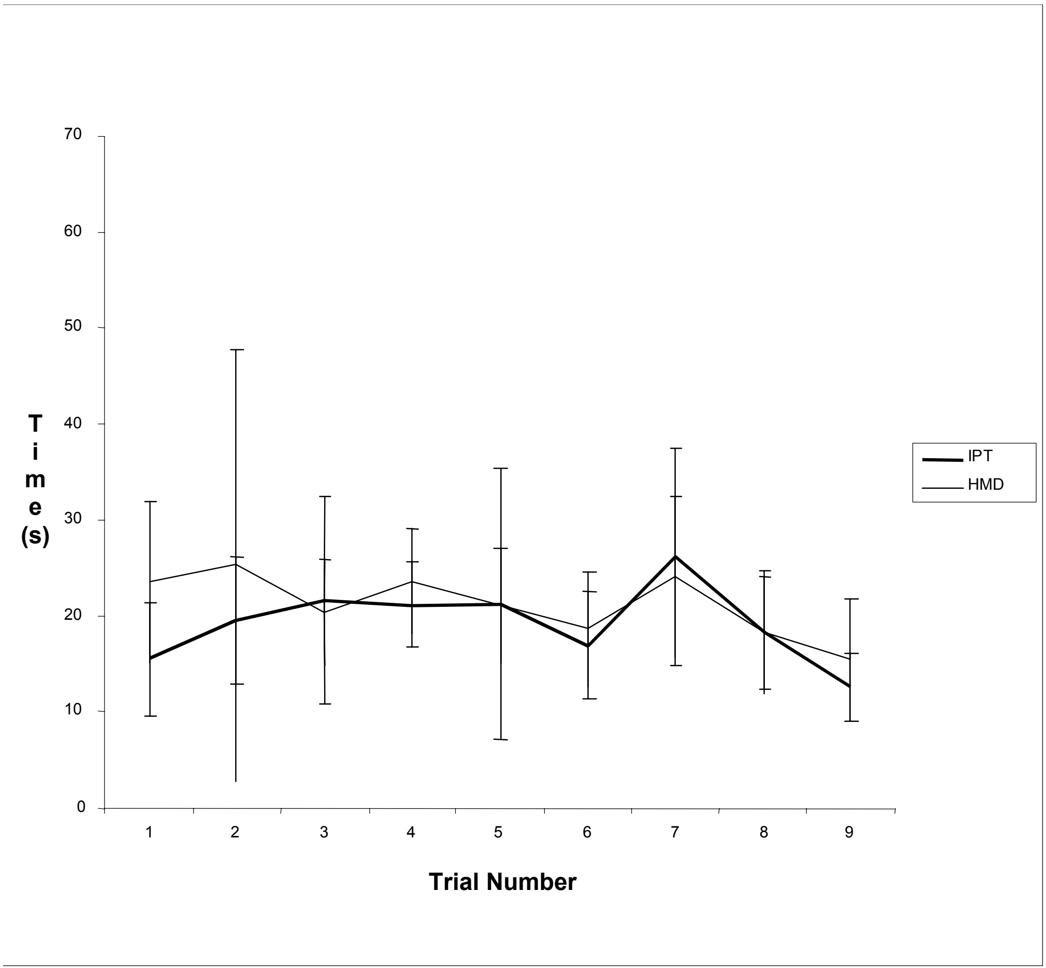

Figure 7 Comparing average trial completion time across display condition for the selection manipulation (SM) task with ray casting

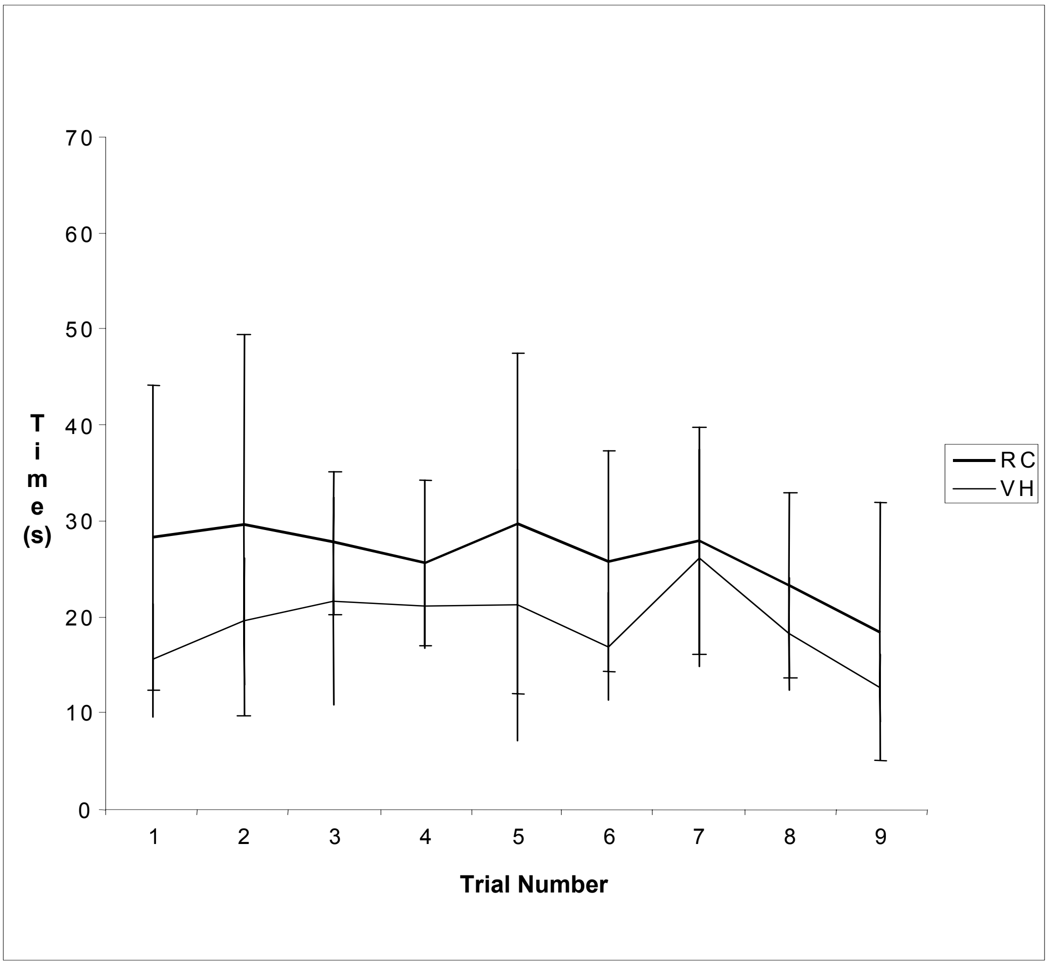

Figure 8 Comparing average trial completion time across interaction metaphor for the selection manipulation (SM) task on IPT

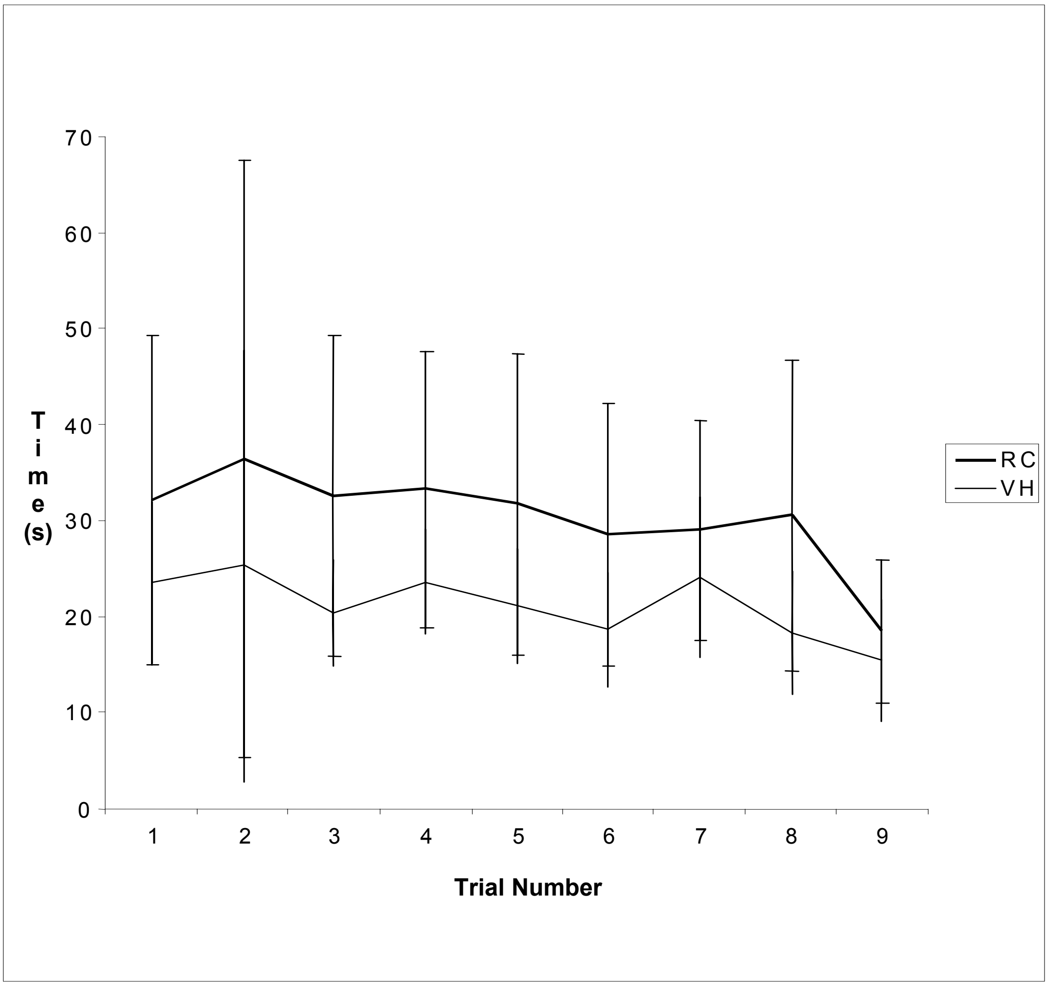

Figure 9 Comparing average trial completion time across interaction metaphor for the selection manipulation (SM) task on HMD

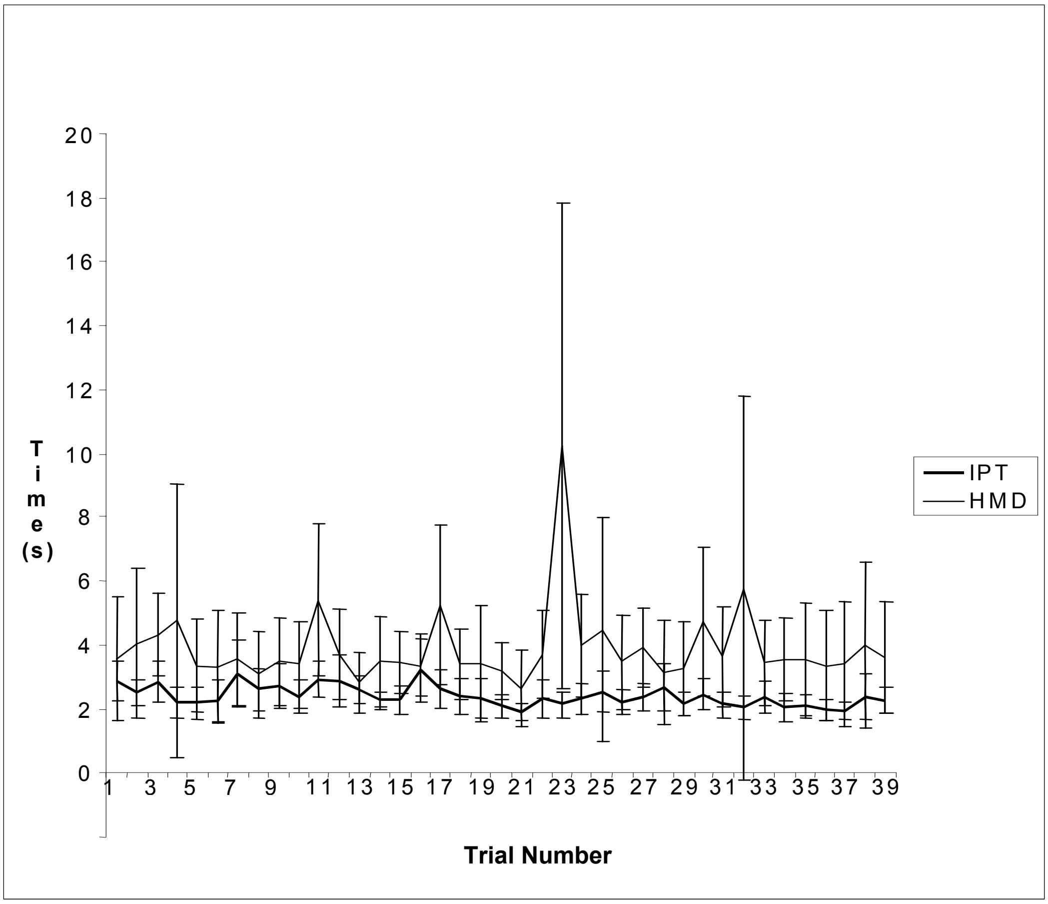

Figure 10 Comparing average trial completion time across display condition for the medium space selection (MSS) task with ray casting