

Impact of Spatial Reference Frames on Human Performance in Virtual Reality User Interfaces

Marc Bernatchez

École Polytechnique de Montréal/Industrial Engineering, Montréal, Canada
Email: marc.bernatchez@polymtl.ca

Jean-Marc Robert

École Polytechnique de Montréal/Industrial Engineering, Montréal, Canada
Email: jean-marc.robert@polymtl.ca

Abstract— The design of virtual reality user interfaces (VRUI) is still an open field of research and development. One category of VRUI is the 3D floating menus that can be manipulated by users in free space. These menus can contain various controls such as buttons, sliders, and text. This article presents an experimental study that aims at testing the impact of five spatial reference frames on human performance with VRUI. Fifteen subjects participated in the study. Wearing a head-mounted display (HMD) and tracking hardware, they were asked to move a slider on a floating menu in a 3D virtual space. The menus were completely virtual such that no physical surface was providing haptic feedback to the user. Main results show that the reference frame has a determinant impact on task execution time, positional accuracy and user preference. We also found a link between the number of degrees of freedom shared between these reference frame and user performance. Results also show that user preference can greatly differ from experimental results obtained by measuring task completion times and positional accuracy.

Index Terms—Virtual reality, user interface, VRUI, 3D floating menus, reference frame, user performance study

I. INTRODUCTION

The challenges in designing and using user interfaces inside virtual reality (VR) systems are numerous. There are very few hardware pointing devices available and they are not yet standardized, the software interaction modes between the user and the VR interface are not well established, and the displays used for presenting information (e.g., CAVE, HMD) are not optimal. This makes it difficult to design 3D environments with good quality virtual reality user interfaces (VRUI). Another problem when using VRUIs is that the interaction occurs in free space, along three axes, namely x, y and z. This increases the complexity of manipulating the user interface (UI) control elements.

This study investigates 3D floating menus which

provide a rich set of UI controls to the user. The floating menus face the same problem of being 3D elements which are floating in free space. Since they are virtual, these menus do not have physical counterparts in the user's world. This creates various manipulation problems. One of them is that the menus do not offer any physical resistance to the user's movement. For example, as the user moves his finger toward the surface of a floating menu, he will not feel anything as his finger goes through that menu surface. This raises the question about our ability to manipulate this type of UI with the difficulty brought by the addition of the third dimension.

These concerns relate to the spatial nature of these floating menus. Thus, we wanted to determine if the way these floating menus are used in 3D space had any impact on the user performance and satisfaction while manipulating the VRUI elements.

The reference frame of a floating menu defines a relative spatial relation with another 3D point in space. We identified two important points in space to relate these menus to: 1) the virtual world origin and 2) the user position in space. These two points represent two opposite references in space as one (the user) is closely related to the other (the world origin). The VRUI can then be seen as a 3D object that must lie in-between these two 3D space references.

As will be seen in upcoming sections, we found that the reference frame used has indeed an impact on the task completion times as well as the positional accuracy. We also found that as the number of degrees of freedom shared by both frames augments, the user performance improves. These findings will provide great guidance for future VRUI developments in order to maximise user performance.

II. PREVIOUS WORK

In 1992, Ellis & Jacoby [1] published the results of a study about menus used in a virtual reality environment.

Based on "A Study on the Impact of Spatial Frames of Reference on Human Performance in Virtual Reality User Interfaces", by Marc Bernatchez, Jean-Marc Robert which appeared in IEEE International Conference on Systems, Man and Cybernetics, 2007, 7-10 Oct. 2007. Montreal, Que., Canada : IEEE. P. 2600-5.. © 2007IEEE.

They conclude that virtual menus are an effective way to interact with a virtual environment. Moreover, they find that visual feedback is an important factor in making these menus effective. They also found that the use of hand gestures should be limited as much as possible as it can reduce user efficiency as the amount of hand gesture patterns augments. This goes along with our principle that VRUI should employ as few input mechanisms as possible. For instance, we believe that the use of buttons on a Wand pointing device should be minimised to a strict minimum.

In 1995, Wloka & Greenfield [2] introduced the "Virtual Tricorder" which uses the concept of mapping. They created a virtual representation of the real input device to which a floating menu is attached inside the virtual environment. This system has the particularity of being highly reconfigurable as the displayed menus can be changed based on the context the system is in. The same physical pointing device can thus be reused in different menus and virtual environments, and for various needs. It also provides a more stable manipulation as the user holds the physical device in his hand and is able to act on the interface through buttons localized on that input device.

In 1996, Mine introduced many important concepts relating to floating menus for virtual reality systems [3]. One of the key concepts he bases his work on is proprioception which allows one to know the relative positions of his body parts without relying on the vision sense. He also exploits interaction methods that involve the two hands. He proposes a floating menu system where the menu is actually attached to the left hand while the right hand is used to interact with controls on the menu surface. The pointing method he uses is based on the "laser ray" approach where a straight line is drawn and oriented given the right hand index finger. The selection is achieved by performing specific finger gestures with the right hand. Finally, he proposed the "tear-off object palette" which is similar to the "World in miniature" (WIM) proposed by [4].

Some researchers were interested in the user performance with these floating menu interfaces. One of them is Coninx [5] who propose a system using conventional 2D window elements such as those found in the X-Windows / Motif exploitation system. Their proposal is that these 2D menus are already well known by the users and are thus assumed to be easier to learn and use in the context of a 3D virtual environment.

In 2000, Kim et al. [6] conducted a study on the impact of menu presentation (the way UI elements are disposed on menus) and positioning with respect to the scene and the user. They compared three positions: 1) static with the scene, 2) static with the user view, and 3) static with a scene object. They also considered various textual menus arrangements such as: pop-up, pull-down, fix stack, basket stack, object pop-up, object pull-down, and

oblique-layered. The authors used different input devices to evaluate the menu conditions, thus making it difficult to conclude on the impact of the menu composition itself on performance.

In 2001, Lindeman investigated the impact of constraints¹ and physical retroactions (with physical surfaces) on interaction efficiency using virtual surfaces (e.g., virtual buttons) [7]. Results show that the use of constraints alone does not improve user performance whereas the use of a physical surface such as the "pen & tablet" [8] does. A limitation of this study is that the authors used a monoscopic head-mounted display (HMD) which may have a significant impact on the user's depth perception. With a virtual surface, depth perception is a key element when pointing and interacting with objects. Using a physical surface, the lack of depth perception can be compensated by tactile feedback.

In 2003, Lemoine et al. identified some problems with floating menus. Located at the user's head height, they cause fatigue over long periods. Item selection from such menus is also problematic due to an inefficient use of hands. They make this interesting comment:

"... as with the integration of the mouse to the PC, we must now find an efficient mean to interact in VE, because we are not anymore in the age of the typewriter, keyboard, and mouse, it is good to innovate and find the best way to interact with the new free space. Perhaps, using 2D space desktop all the time is maybe an inhibitory parameter that prevents us to create good 3D interfaces." [9]

In 2003, Larimer and Bowman [10] tested user performance with three types of VR floating menu tasks: 1) opening, 2) closing and 3) moving. Results are preliminary and based only on five subjects. They show that performance widely varies from subject to subject, 2) the floating menus are usable yet not as efficient as their 2D counterparts, and 3) the Wand (a pointing device) imposes precision limitations which makes the closing of menus difficult.

In 2005, Kohli and Whitton [11] presented a survey on the effects of the use of the non-dominant hand to manipulate elements of interface on a floating menu. The floating menu is held by the non-dominant hand whereas the dominant hand points the elements of interface. Their survey has some limitations. One of them is that only 8 subjects participated in their survey. This is insufficient to draw reliable conclusions. Another limitation is that they did not make a comparison with a second interaction

¹ Restricting UI element behaviours following certain rules so that it is easier to manipulate for the user. For example, disallowing a cursor button to move or slide following given axes so that it can only move or slide following the intended (useful) direction. Constraining a UI element is equivalent to locking behaviours that would rather impede the user than help him achieving a given task. See [13,14]

technique acting as reference. It makes their findings hazardous and difficult to interpret.

In 2006, Dachsel and Hübner [12] proposed a taxonomy of interfaces for "control" menus which does a relatively complete overview of interfaces based on floating menus. They noted that the majority of the interfaces they examined were in the "simple menu" category, i.e. menus generally using less than 8 elements that can be displayed over an either short or long period of time. The literature review that they present is interesting, but the taxonomy itself doesn't bring any real answers on the position of each of these interfaces with regard to the others. The first level of division of their taxonomy is based on four possibilities: the temporary options menus, the simple menus, the systems menus and the hierarchical menus. One problem with this taxonomy is that the choice of the division criteria is relative to a very particular need that is not necessarily the one required by a given application. Their taxonomy is not therefore sufficiently generalizable in our opinion.

As it can be seen from this, most research efforts were focused on developing new user interface techniques and visual widgets. On the other hand, the number of studies on the user performance of such techniques and widgets remains scarce. The experiment we describe here is the first of three studies we did in that direction.

This experiment aims to test the impact of different spatial frames of reference on user performance and satisfaction in VRUI. In the next two sections, we present the frames of reference and the methodology of our experiment. We test different floating menu configurations which are attached to various degrees to the user's body.

III. THE FIVE REFERENCE FRAMES

We considered five different reference frames to define the dynamic positioning of a floating menu in relation to the user's body. In order to establish which reference frames to use in our experience, we considered the link between the referential (the position and orientation in space) of the user and the floating menu referential.

A convenient method to classify these links is to consider the number of degrees of freedom (DOF) that are either shared or common for two referentials. For example, if the floating menu moves following the position but not the orientation according to the user's movements, then we will say that these two referentials share three degrees of freedom (translation on x, y and z axes). The changes of position or orientation that are transmitted from a referential to the other are said common or shared. The choice of the five reference frames that were retained for this experiment has also been influenced by practical considerations. Thus, we considered reference frames that represent real world applications cases. For example, we could have considered reference frames where only one or two degrees of freedom were common. It could have taken the form of a stationary menu in relation to the scene according to two axes, connected to the user's head. The

usefulness of such cases is marginal in a real use context because it is not useful for a floating menu to follow the vertical movements of the user's head if the user is far from that menu, and thus unable to clearly see or use the menu. We thus excluded these cases of our survey. Our final choice of five reference frames has been made in order to obtain a gradation of the number of degrees of freedom as indicated in table 1. It allows us to perform a methodical and systematic assessment of the reference frames according to the number of shared degrees of freedom.

Table 1 presents a summary of every reference frame features as well as the number of degrees of freedom that are used. We ordered the reference frames in increasing order following the number of degrees of freedom with regard to the user's body position and orientation. This allows one to see if there is a correlation between the degrees of freedom controlled by a given reference frame and the observed user performance. We chose these reference frames so as to cover a set of cases that include the two extremes regarding the relation between the user and the interface referential that must be manipulated in the VR environment.

TABLE 1
EXPERIMENTAL CASES

| Frame # | Reference frame | Description | DOF link |
|---------|---|---|----------|
| 1 | Static with scene | The menu does not follow any user's movements. | 0 |
| 2 | Static following the position of the user's body | The menu follows the user in position only. | 3 |
| 3 | Static following the position and one axis of rotation (heading) of the user's body | The menu follows the user's position and turns to remain facing the user. | 4 |
| 4 | Static following the user's left hand | The menu is seen as being attached to the user's left hand. | 6 |
| 5 | Static following the user's head | The menu is following user's gaze direction following a head-up-display (HUD) manner. | 6 |

IV. METHODOLOGY

A. VR Test Platform

The VR test platform we developed incorporates multiple different floating menu configurations. It shows

a natural scene that users are familiar with which consists of a house interior. This environment provides scene occlusion conditions and color contrasts which provide a typical and plausible virtual environment context. The VR test platform presents two floating menus to the user. One is static with regard to the scene and allows the user to control the experiment as well as to see textual information about the current state of the experiment (see figure 1). Two buttons located at the bottom left and right of this menu allow the user to trigger the beginning of a new trial (labelled “D butter” on the figure) or to perform a practice trial (labelled “Pratiquer” on the figure), the only difference with the practice trial being that no actual data are recorded.

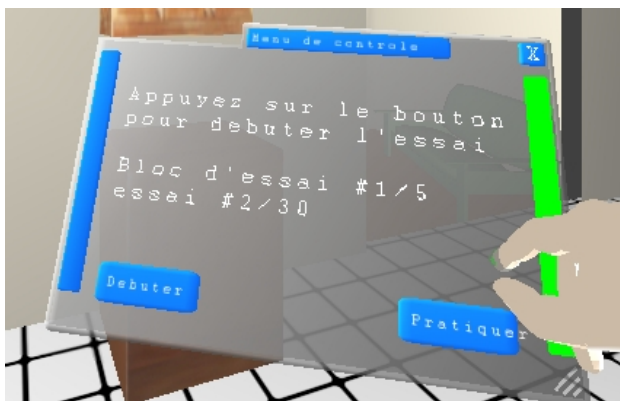


Figure 1. Experiment control panel²

The other floating menu provides the actual experiment UI elements on which the user performs the experimental task.

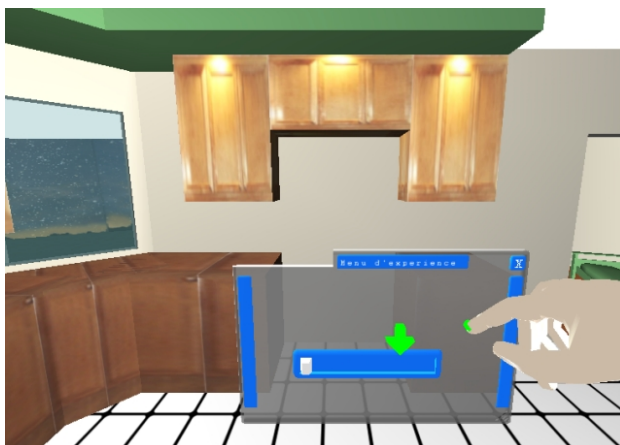


Figure 2. View of the VRUI test platform and main experiment floating menu

It contains the interface elements on which we measured the subject's performance and satisfaction. This menu includes a linear scrollbar with a cursor that can slide, following its longest axis. The length of the

scrollbar on the sliding axis is of 15.6 cm by 3.3 cm in height. The dimensions of the cursor are 1.2 cm (horizontally) by 1.9 cm (vertically). A green arrow is fixed above the scrollbar to indicate the release point target position for the cursor. The dimensions of the arrow are of 5.3 cm in width by 5.5 cm in height. The thickness of the interface elements is of 1 cm. The menu itself measures 30 cm wide by 20 cm in height by 1 cm in thickness. These dimensions are on scale so that the dimensions perceived by the user of the VR system were as indicated above. For example, when the user did a horizontal movement, starting from the left extremity going to the right extremity of the menu, his hand did a physical movement of 30 cm in the real world, as perceived in the HMD. The figure 2 shows the menu with its components.

The hardware equipment used was composed of the following components: a Pentium 4 PC running the Microsoft Windows XP operating system, a magnetic tracking system and a head-mounted display (HMD) helmet. The scene rendering was achieved using an NVIDIA GeForce 5500 3D accelerator card. This card features two video outputs allowing us to simultaneously connect a LCD monitor and the HMD to the system. This allowed the experiment coordinator to monitor the user's manipulation inside the virtual world in real-time. The tracking system that we used was the Wintracker from VR-Space, Inc. which uses electro-magnetic signals to track up to three points in space following six degrees of liberty (6 DOF). The tracking volume is of approximately 1.5 meters on each axis³. The HMD used was the I-glasses PC/SVGA Pro 3D from i-O Display Systems. It has a resolution of 800x600 color pixels for each eye and a field of view of 25 degrees (diagonal). This device can show a stereoscopic scene using separate left / right eye image pairs.

For the software of the platform, we used two public domain libraries. The first is VRJuggler 1.0.7 [11] which provides functionalities to interface with various VR devices. The second library is OpenSceneGraph 0.9.6 [12] which provides functionalities to describe a virtual scene following a structured hierarchy. It also provides the capability to load 3D geometry files in order to populate the scene hierarchy.

B. Task

The subject was asked to move a cursor on a horizontal slider from its starting position to a destination indicated by an arrow located on the floating menu surface. The starting position of the cursor was always the same, on the left extremity of the slider. The green arrow position was stationary on the floating menu, for the whole experience duration. The travel distance was therefore the same from a trial to another. Figure 3 shows the manipulation to perform.

² We can read on the panel: “Push on the button to start trials. Trial block #1/5 – Trial #2/30 – Begin [left button] Practice [right button]”

³ A detailed review of the Wintracker is available online at http://vresources.org/Wintracker_Review.html

The instructions were to perform the task as rapidly as possible, while trying to release the cursor as precisely as possible under the tip of the arrow. The cursor was put back at its starting point at the extreme left of the slider bar at the beginning of each new trial. The subject was not imposed a time limit to perform the task. The end of the trial was determined by the cursor position as it was released. The trial timing clock was stopped only when the cursor was released inside a predefined zone defining a tolerance region around the release point marked by the arrow. This region covers 8% of the maximal cursor displacement on the slider bar, on each side from the green arrow center point. Since the total length of the slider bar is 12.8 cm, the allowed release zone is 1.08 cm on each side of the release center point. Since the menus are presented following a 1:1 scale, the subject was allowed to physically release the cursor inside an area covering 2.15 cm that was centered on the green arrow like shown on figure 2.

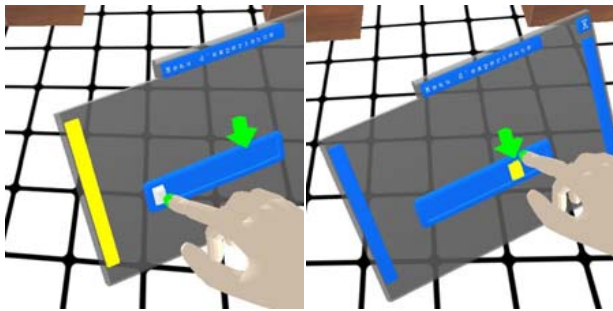


Figure 3. The experimental task. Moving a cursor on a horizontal slider and releasing it under the tip of the arrow

We defined a release zone that was narrow enough to be a challenge for the subjects and force them to cover the same distances with the cursor..

We did not want that a subject could obtain a trial time shorter by releasing the cursor prematurely, and thus being less precise in his movement, than another subject which would be trying to be more precise performing the task. While imposing this release zone window, we wanted to ensure that it would be difficult to cheat so to speak. Preliminary tests showed that the prescribed zone was adequate in this regard and offered a good compromise between precision and manipulation speed. The release zone was found sufficiently large to allow to perform the task in reasonable delays (our target average trial time was to be lower than 10 seconds) while considering the constraints imposed by the VR equipment used (tracking noise and pointing precision) as well as the dexterity of the subjects. Finally, it is to be noted that the subjects were not informed about the physical dimensions of this release zone. They were only informed of the existence of this zone.

C. Subjects

Fifteen subjects, composed of 11 men and 4 women, took part in the experiment. The average age was 27 years old, varying between 18 and 60 years. The selection criteria used were: being right handed and not being

color-blind. The subjects had little or no past experience using VR systems.

The required prerequisite to be right-handed was imposed for technical considerations. Our menus have indeed been conceived for a use by right-handers. For the purpose of our tests, we were not interested by the differences between right-handed and left-handed users. The inclusion of these in the tests would have required us to add support for this in our test platform in order to be able to manage this case. The second prerequisite regarding the color vision essentially comes from the fact that the virtual environment uses color changes for some interface elements in our tests. It is therefore necessary for an individual to well discern these color nuances in order to be able to do the various tests. We verified this condition at the beginning of each experiment with the subject using the helmet that presented interface elements with various colors.

The subjects had no or very little previous experiences in the use of virtual reality systems. Only one subject was present at a time in our experiment laboratory. The subjects were asked to read and sign a consent form before beginning the experience.

D. Procedure

At the beginning of the experiment, each subject had to complete a brief questionnaire about his age, gender and past experience with computer systems as well as VR systems. Next, a brief presentation of the VR test platform was made. This presentation was in two parts.

Firstly, the experiment coordinator would put on the VR equipment and show to the subject how to proceed to accomplish the various manipulations inside the virtual environment. He would also give recommendations on how to interact with the interface. He explained to the subject how to move his hand at proximity of the floating menus and how to make the final approach toward the UI controls with his hand, by passing in front of the menu. We told the subject to hold his arm alongside his body after he had pushed the "D butter" button on the control menu and to wait the appearance of the second floating menu. We asked the subject to approach the floating menu from the front with his finger (the pointer) such that he would systematically pass in front of the menu. This ensured a consistency in the movements made by the subjects before starting the chronometer. This introduction lasted around 5 minutes.

Secondly, the subject put on the VR equipment and practiced using the VR test platform for a 5 minute period. We determined, by preliminary tests, that this practice period was adequate to allow the subject to acquire a minimal comfort and ease with the various UI manipulations while remaining able to detect the learning curve effect in our measures. No measures were taken during the practice period.

Once the practice period was over, the subject took a short pause during which we informed him that the next trials would be recorded. The experiment period lasted 45 minutes on average, excluding the introduction and

practice periods. The trials were grouped in 5 blocks of 30 trials each, for a total of 150 trials per subject. Each of these blocks represented one of the five different frames of reference that are being compared in this experiment. The order of presentation of each five cases was randomized for each subject to minimize the learning effect on the data.

At the end of each block, we asked the subject to indicate his preferred frame compared to the previous ones. Any other comments were noted for further analysis.

E. Experiment Variables

The main independent variable is the reference frame. The subject gender and the tracking noise are secondary independent variables which were considered due to the impact they could have on results.

The main tracking noise sources were electromagnetic interferences perceived by the tracking system. The environment in which our laboratory was situated contained multiple potential interference sources. The wall, ceiling and floor structures are made of armed concrete which includes metallic grids. These grids act as electromagnetic sources as they reflect electromagnetic fields present in their surroundings (in this case, the signal emitted by the tracking system). The tracking system sensors were receiving parasite echoes of the main tracker signal plus the main signal itself.

The second interference source was the 120 Volts electric power lines that are incorporated in the walls of the room. The current going through these lines generate low-frequency electromagnetic fields for which the amplitude can be substantial in short ranges.

The dependent variables are the trial time, the positioning error, and the subject's satisfaction (preferred reference frame). For the trial time, we actually measured the cursor selection (picking) and the cursor manipulation (moving) sub-times. These variables were compiled and recorded in a computer file at the end of each trial.

The trial time includes the time interval beginning when the virtual hand enters into the active zone of the floating menu and ending when the cursor is released. This zone forms an invisible cubic volume around the menu. This volume had a fixed size from a trial to the next. The volume was 30 x 20 x 7.5 cm (width / height / depth). The base of the volume lies on the surface plane of the floating menu and extends 7.5 cm in front of it. As the travel distance is 7.5 cm before reaching the cursor, we could eliminate all measured time intervals related to the virtual hand travel before it gets at proximity of the menu. The initial position of the subject's hand in space was thus not influencing the measured times.

The positioning error was measured by computing the distance between the ideal release point (at the tip of the green arrow on the floating menu) and the actual point where the subject released the sliding cursor.

The user satisfaction was measured through a short questionnaire at the end of the session. Each subject was asked to indicate his preferred reference frame, his

fatigue level and any physiological discomforts he could feel.

V. RESULTS

A. Descriptive statistics

Figure 4 shows that the reference frames did have different impacts on the subject's performance with regard to trial times. Frame 3 yields the best average trial time (5.26 sec) while frame 4 yields the worst average trial time (6.08 sec), making the largest difference in trial times to be 0.82 seconds (13.5%).

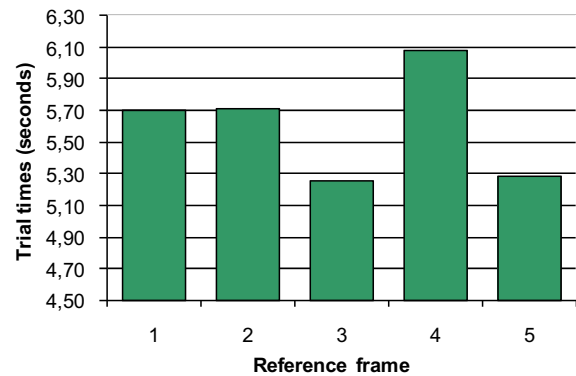


Figure 4. Average trial times for each reference frame

Thus, the reference frame choice can have a significant impact on user performance over extended immersion periods. From these numbers, a user could save as much as 8 minutes over a 60 minutes immersion period, allowing him to achieve the task in 52 minutes instead. There are also significant average time differences on secondary variables of 19% for the gender and 12% for the tracking noise.

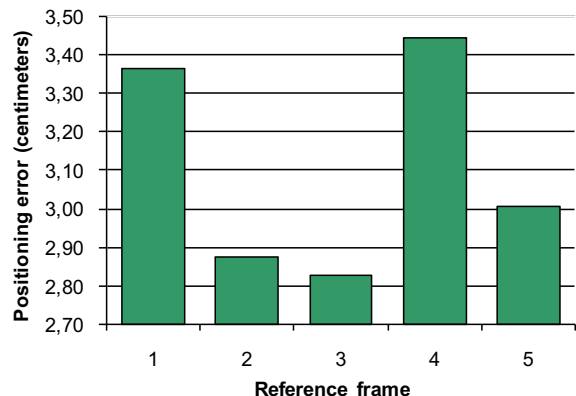


Figure 5. Average positioning error for each reference frame

Figure 5 gives the performances of each reference frame regarding the positional error, that is, the distance difference between the cursor release point and the marked release position. Figure 4 shows that frame 3 had the lowest average positioning errors (2.83 cm), whereas frame 4 had the largest one (3.44 cm). The positional

error variation is thus 17.7% between the worst and the best cases.

The second independent variable considered is the gender of the subjects. Considering the low number of women subjects, the results of this section are given for information only and serve to nuance the possible differences that could exist regarding genders. Further studies would be necessary to establish with certainty if there is a meaningful effect related to the gender.

In the case of the independent variable representing the gender of the subject, we note a concordance of the performance results with the relative dependent variables related to time. The men are on average faster than women. The figure 6 gives a graphical representation of the situation.

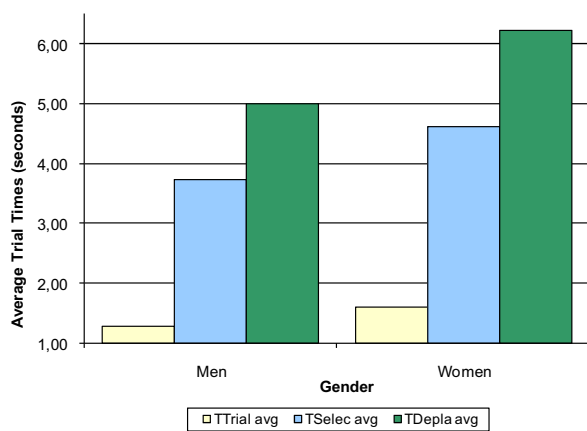


Figure 6. Total average trial time for every gender

If we look at the positioning error, we see that women are more effective than men this time. The manipulation precision is greater for women. It therefore seems that there is a pendulum effect between the manipulation time and the precision according to the gender.

We see in figure 7 that the women are slower but more precise than men. On the other hand, we cannot draw any definitive conclusions related to the gender due to the limited number of women subjects who took part to our experience.

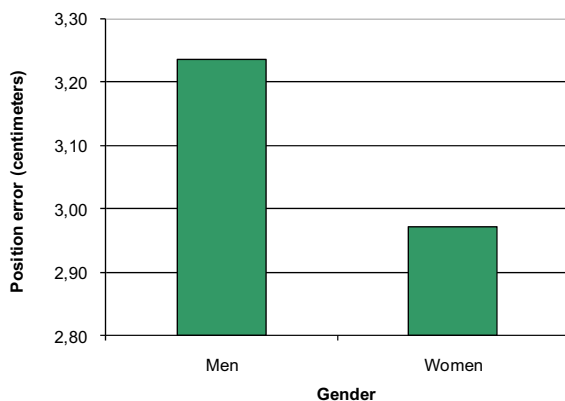


Figure 7. Average positioning error for every gender

In terms of relative differences, we note a 19% difference between the genders regarding the total average times, which is considerable. It therefore seems an element to investigate further in future experiences regarding this topic.

The figure 8 gives the results regarding the impact of the noise introduced by the positions sensors on the measured times. There is a perfect concordance between the variable "sensor noise" and the observed averages. We note that the sensor noise caused an increase of the measured trial times in the order of 14% for all the subjects. The sensor noise is therefore a factor that had a non negligible influence on the results of this experience. It showed to be harmful on the user's performances at the time of the use of a VR interface.

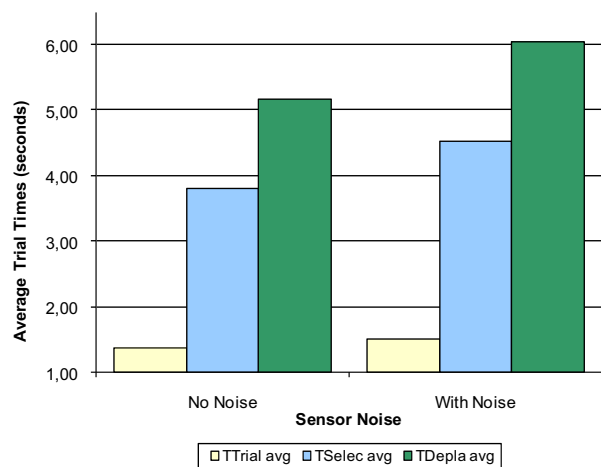


Figure 8. Total average trial time with regard to sensor noise

The noise had an impact on the positioning precision in the order of 8%. According to these results, the impact on the precision is less important than on the manipulation times. A possible interpretation is that the noise causes an hesitation of the subject at the time of the start of a movement. Based on the previous data, this hesitation seems to dominate on the positioning error.

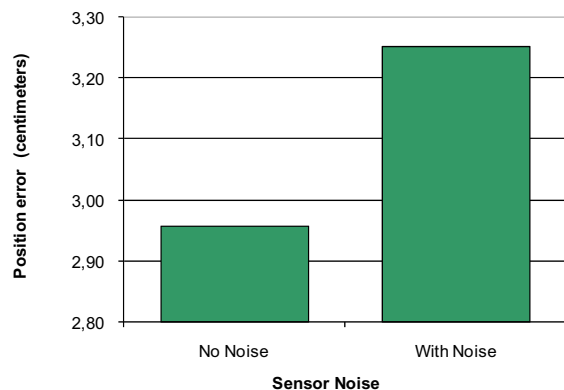


Figure 9. Average positioning error related to sensor noise

These observations demonstrate us that there was a non negligible effect on our results that is due to technical

limitations related to the equipment used. The figure 9 shows the positioning error according to the sensor noise.

B. ANOVA analyses

We performed ANOVA analyses on our measurements and found statistical significance for the gender factor in the average trial time case with $p=0.0122$ and $F=6.72$, but since the number of women is low ($N=4$), we cannot conclude in that regard. We did find statistical significance for the reference frame factor when measuring the minimum cursor selection time with $p=0.0089$ and $F= 3.77$. In that case, the gender is also present as a dominant factor ($p= 0.0263$ and $F= 5.22$) even though we cannot conclude regarding the gender factor.

VI. DISCUSSION

A. Best reference frame

Not considering the gender and the sensor noise, the reference frame 3 (stationary with regard to the position and the direction (1 axis) of the user's body) gave the best performance results. This reference frame shares four degrees of freedom with the referential of the user's body. The reference frame 5 follows closely in terms of performance. It shares six degrees of freedom with the referential of the user's body. We will come back on the relation between the number of degrees of freedom and the performances later-on.

The reference frames 3 and 5 are those that offer the best average trial times. The reference frame 4, relative to the non dominant hand, is the one that had the worse trial times. Figure 10 shows these results. Several reasons can explain this situation. In the first place, the reference frame 4 uses three position sensors rather than two for the four other reference frames. It implies a supplementary source of sensor noise, creating more instability on the floating menu position. It will be therefore more difficult for the user to do the manipulations because he will need to compensate for the movements induced by each of these sensors.

Indeed, every position sensor is actually a set of three directional antennas that captures radiofrequency signals transmitted by the emitter of the system. The noise comes from various sources of electromagnetic fields such as high tensions lines. The noise perceived by the position sensors is equivalent to random white noise having a certain amplitude. Each of the sensors used see this noise that it then relays to the tracking system combined with the useful signal. The noise disrupts the perceived positions in a randomly manner, and this, with different errors in time for every sensor. In other terms, if the noise of the sensor #1 is such that the position is off by x millimetres on the left, the sensor #2 can exhibit an error of y millimetres upwards or on any other axis, in a random manner. When the positions returned by these sensors are combined together, these errors combine to give a composite displacement z that is the vectorial sum of all present errors. The errors add themselves and therefore becomes more important according to the

number of sensors (noise sources) used. It directly affects the movement coordination of the user with regard to the floating menu because the noise introduced in the system varies quickly in time, causing fast and perceptible oscillating movements of the virtual objects such as floating menus.

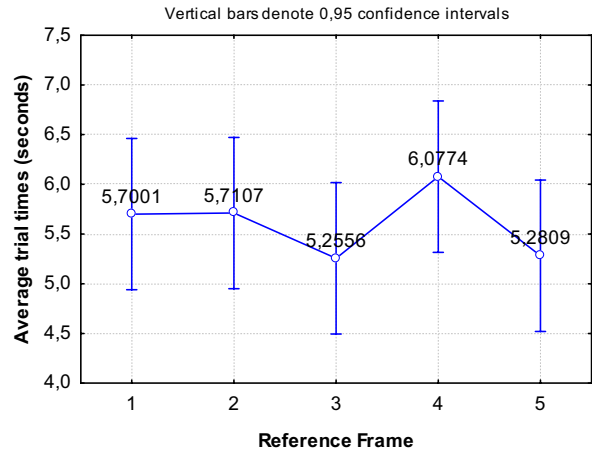


Figure 10. Total average trial time for each reference frame

B. Shared degrees of freedom

What is more interesting is to see the resulting evolution of the minimal cursor selection times for each reference frame on figure 11. Notice that the order in which they are given on the graph X axis is such that they are in growing order relative to the number of shared degrees of liberty, as mentioned earlier in table 1. What we can see on the graph is that the times improve as the VRUI reference frame shares more degrees of liberty with the user's body reference frame.

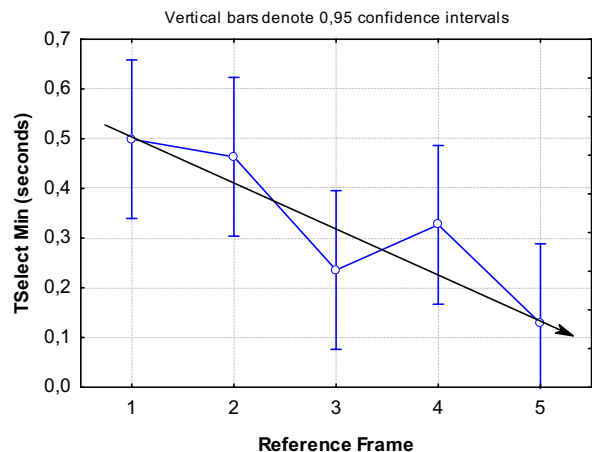


Figure 11. Minimal cursor selection times for each reference frame

The reference frame #1 is characterized by the fact that it doesn't depend on any position sensor. Conversely, as it is stationary with regards to the scene (to the virtual world), it becomes mobile with regards to the user. The user's movement results in an opposite movement of the floating menu as seen from the reference frame of the user. For example, when the user moves toward the menu, from the user's viewpoint, it is the menu that goes

toward the user following a trajectory of opposite direction. Likewise, if the user lowers toward the ground, the menu will seem to go up to the eyes of the user. It makes it more difficult for the user to coordinate his movements because the menu is seen like being mobile to the user's point of view.

The figure 12 shows the results for the positioning error. We observe that the reference frame #3 gives the best performances. The reference frame #5 had poorer performance than the four other frames with regards to the minimum trial times. Hence, two of the five reference frames had performances lower than the others. These are the reference frames #1 and #4.

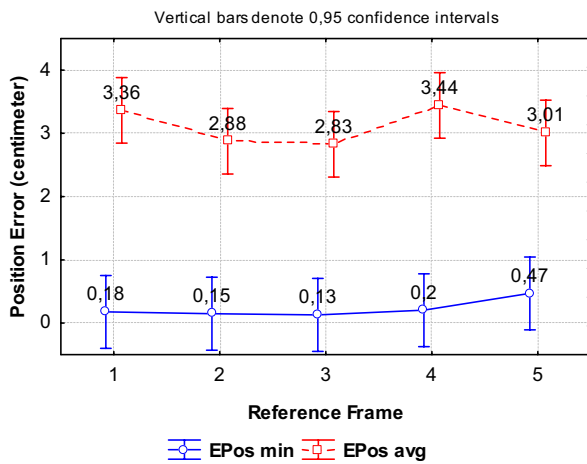


Figure 12. Positioning error for each reference frame

C. Learning curves

Figure 13 gives the learning curve of the subjects for each successive trial. The observed learning is accentuated in the beginning of a trial block to later reach a minimum toward the 24th trial. Then Follows a light performances deterioration showing what we can interpret as the apparition of fatigue signs on the last 5 or 6 trials. This shows that the number of trials used in this experiment is appropriated. It also denotes that the use of the VR system requires a certain training period before the user becomes comfortable with the manipulations to perform.

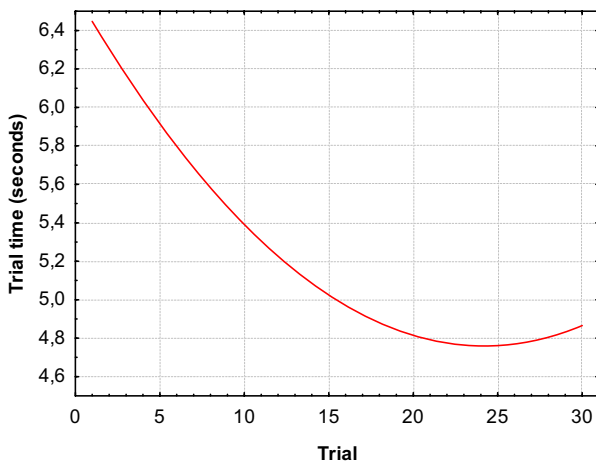


Figure 13. Learning curve with regard to trial times

Figure 14 gives the learning curve regarding the positioning error according to the trials performed. Here, the learning is faster and reach a minimum toward the 16th trial. On the other hand, a quick performance deterioration occurs after the 16th trial, which we interpret as the effect of fatigue. This fatigue appears earlier for manipulations errors than for the manipulation time.

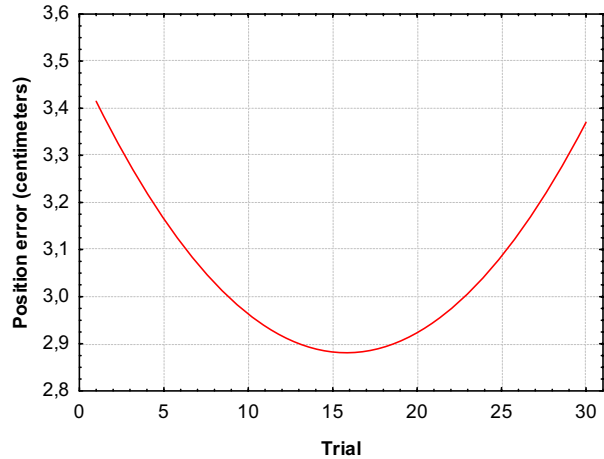


Figure 14. Learning curve with regard to positioning errors

D. Differences attributable to the gender

The gender is a factor that influences the user's global performances strongly as we noted it previously in the results of the ANOVA. Because of the low number of women subjects that were available at the time of our experience, we cannot draw any strong conclusions on the effect of the gender on the user's performances. The following discussion only allows to evoke a possible effect of the gender factor on the performances and must be considered therefore with prudence.

In figure 15, we observe that the women are in general slower than the men. This performance gap is on average of more than one second per trial. This difference of performance doesn't seem to be specific to a particular task such as to manipulate the cursor or to select it as the two curves have similar slopes. It confirms that there is a difference based on the user's gender.

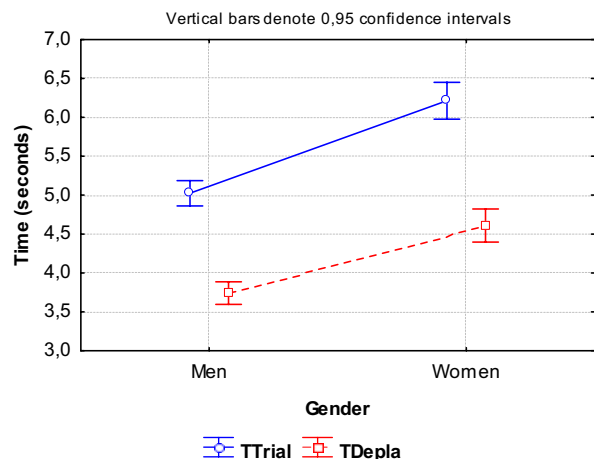


Figure 15. Trial times with regard to gender

On the other hand, if we look at the precision of manipulation of the HMI (human-machine interface), the women are more effective than men. Another point that we note looking at figure 16 is that the variability is bigger for women than men. The women are therefore in general more precise but have less constant or steady performances. The lower number women subjects could also explain this.

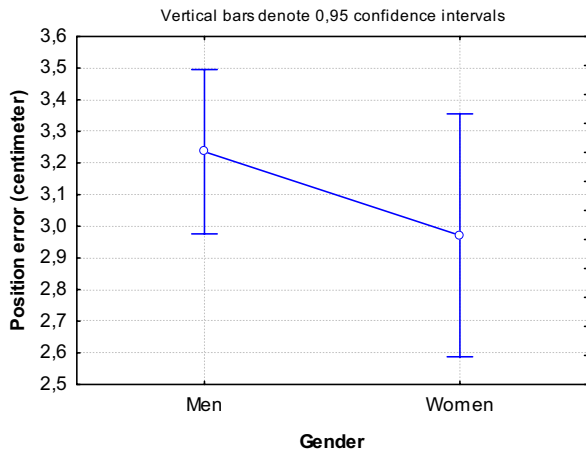


Figure 16. Positioning error with regard to gender

E. Qualitative results

At the end of the experiment, each subject answered a few questions on his preference for a reference frame. Figure 17 shows the results on user’s preference regarding the reference frame used.

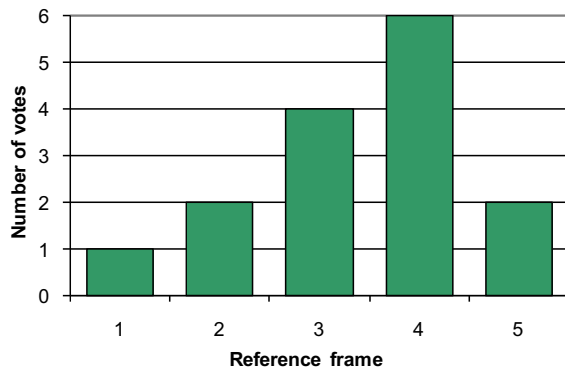


Figure 17. Subjects’ preferred reference frame

The reference frame 4, where the floating menu is attached to the user’s left hand (not dominant), was preferred by 6 subjects out of 15. Surprisingly enough, the subjects preferred the worst case based on figure 3 and 4 results. This can be attributed to the importance of the non dominant hand contribution in the context of a two-handed user interaction [13,14]. The use of both hands during a user interaction allows one to benefit from better feedback provided by proprioception that the majority of the population possess. Moreover, the two hand approach allows to associate the menu reference frame to the left hand. By way of the proprioceptive feedback, the left hand reference frame can be directly

integrated by the user’s brain and used in conjunction with the right hand reference frame which is also directly linked to the brain through proprioceptive cues. Thus, the difference between interacting with a menu attached to the user body compared to a menu using an external reference frame is that in the first case, the reference frame information is directly available to the user’s brain while in the second case, the brain must deduce that information from other senses such as the sight of the menu in space. Why this reference frame performed so poorly is probably attributable to the fact that, for this case, all three tracker sensors were used thus inducing tracking errors three times rather than two times for the other cases. Each tracker is a small antenna that receives radio signals. Doing so, it inadvertently picks-up RF noise from the surroundings which translate into a tracking error noise. The error of each tracking sensor adds up such that the more points are tracked in space, the greater the error. For instance, using the frame 4, the left hand has the menu attached to it. Due to tracking error, the menu position will jitter. Since the right hand is also tracked, its position also jitters, following similar but not identical noise error in time. Thus, both the menu and the pointing right hand have an error that combines to create an overall greater positional error.

Subjects were also asked to tell their impressions during the experiment. The following table gives a compilation of the comments given by subjects as they made their manipulations. The “Count” column gives the number of subjects who expressed a given comment.

TABLE 2
SUBJECTS’ COMMENTS COMPILATION

| Comment | Count |
|---|-------|
| Menus are not on scale / not aligned | 3 |
| Controlling the menu placement seems an important factor | 3 |
| Releasing the cursor is difficult | 2 |
| I like the menu transparency | 1 |
| There is a contrast problem in the scene using the HMD | 1 |
| A hand shadow would be nice | 1 |
| The virtual hand obstructs the view on the cursor | 1 |
| The hand-to-head distance seems to be an important factor | 1 |

The three most important comments made are: 1) that scaling mismatches, even small, between the virtual scene and the user’s physical body parts are a source of confusion that is easily noticeable; 2) that it is advisable to allow the user to manually adjust the positioning of the VRUI floating menus (in our experiment, this restriction was deliberate in order to isolate other variables); and 3) that releasing the cursor was not necessarily an easy task

the way it was implemented in this experiment. The first two comments are linked to the floating menu positioning as they affect the user’s position perception of the menus in space. The third comment is more linked to technical limitations with regard to pointing precision of the equipment that was used.

Data on the fatigue level felt by the subjects during the experience have been collected in order to determine if the use of the VR test platform and floating menus caused unacceptable fatigue. The results of table 3 indicate that the use of the VR system didn't cause strong or high fatigue levels during the experience.

TABLE 3
FATIGUE LEVEL REPORTED BY SUBJECTS

| Fatigue level | Number of votes | % |
|---------------|-----------------|--------|
| 0 - none | 1 | 6,67% |
| 1 - light | 7 | 46,67% |
| 2 - average | 7 | 46,67% |
| 3 - high | 0 | 0,00% |
| 4 - very high | 0 | 0,00% |

We also asked the subjects to indicate us the sensations of discomfort they perceive (if any) at the end of the experiment session. Table 4 shows the list of symptoms reported by the subjects.

The main symptom reported by the subjects is an ocular fatigue that appears as a blurred vision and that we attribute to the weak resolution as well as the low field of view of the helmet used. The neck tensions that are reported can be explained by the fact that the weak field of view of the helmet forces the user to turn the head more rather than to orient his eyes to look around. Another reason of this neck pain is probably related to the fact that the reference frame 5 imposes an increased stability of the head in order to be able to manipulate the stationary HMI with regard to the head. The back pain can be attributable, according to us, to the fact that we asked the subjects to remain standing all along the experience, during near an hour. An interesting fact to note is that the subjects didn't complain about fatigue during the experience, tending to confirm the observations made by other researchers that the use of a simulation presented in a helmet permits to divert the points of interest and concentration of an individual and to decrease their sensitivity to certain external stimuli causing the pain. Finally, the arm fatigue is explained by the fact that the subjects had to constantly point buttons on the HMI using their finger. The user arm was therefore often maintained in an horizontal position, which implies a constant muscular and maintained effort.

TABLE 4
SYMPTOMS REPORTED BY SUBJECTS

| Symptom | Occurrences |
|----------------|-------------|
| Blurred vision | 4 |
| Neck tension | 2 |
| Back pain | 2 |
| Arm fatigue | 2 |

VII. DISCUSSION

A. *Optimal reference frame*

According to the results, the reference frame 3 comes out as being the best regarding the user performance. The reference frame 4 is distinctively less effective, which can be explained by the use of an additional position sensor as well as the use of two hands, both in movement. These factors add to the instability and imprecision with regard to the manipulations of the HMI by the user.

B. *Precision vs. manipulation times*

Based on the results, the positioning error is more affected than the manipulation time by the reference frame used. As one of the main senses used by the subject is the vision, we can expect that there will be a relation between the positioning error and the quantity of the perceived visual information. We refer here to the display resolution, that is to say, to the quantity of display pixels elements used by the display device. The immersive helmet used in this experiment has a resolution of 800 x 600 color elements (pixels) displayed on a field of view of 26 degrees diagonally. By calculations, we obtain the angular resolution, more useful to evaluate the resolution perceived by the user. We refer to the perceived resolution because it differs from the resolution of the display device and is determined by the characteristics of the optic components used in the helmet. The same number of display elements, 800 x 600 in the current case, can allow to read the text or not inside a helmet based on the resolution of the miniature screens used and the field of view prescribed by the optic components (the lenses and mirrors) used. To allow an easier evaluation of the resolution as perceived by the user, we can therefore use the angular resolution notion. It is based on the size of a scene element with regard to the underlying angle of this element as seen by the user's eye. The angular resolution is calculates as:

$$\text{Horizontal angular resolution (arc minutes/pixel)} = \frac{1 \text{ min of arc}}{1/60 \text{ degree}} \cdot \frac{\text{Field of view}}{\text{Horizontal resolution (pixels)}} \quad (\text{Eq. 1})$$

In the present case:

$$\frac{1 \text{ min of arc}}{1/60 \text{ degree}} \cdot \frac{26 \text{ degrees}}{800 \text{ pixels}} = 1.95 \text{ min of arc/pixel} \quad (\text{Eq. 2})$$

The human eye has an average visual acuity of 1 arc minute for the smallest object that can be discerned to the

naked eye. The immersive helmet used in this experiment therefore displays a scene whose perceived resolution is close to two times inferior to the visual acuity of the subject. This can cause a loss of manipulation accuracy due to the resolution loss of the visual feedback on which the subject relies.

The angular resolution of the display device (an immersive helmet in this case) is constant, no matter the position of the virtual objects in the scene. On the other hand, the farther an object is from the observer, the fewer pixels are available to represent it visually. Therefore, for the reference frames where the menus appear close to the user's body, the resolution, and by extent, the quantity and the quality of the feedback will be superior. Indeed, some reference frames impose a stationary and determined distance between the floating menu and the user whereas other reference frames allow a variable user / floating menu distance.

C. Considerations relative to floating menu positioning with regard to the user

In this experiment we tried to limit as much as possible the effects on the measures related to the floating menus positioning with respect to the position of the user's body. It is for this reason that we introduced the notion of a proximity zone on the floating menu. It allowed us to determine a precise moment where the measurement process began, according to a distance known and fixed from a trial to the other with regard to the surface of the floating menus.

This being said, several factors related to the menu position can have had an influence on the measures. One of these factors is the menu surface orientation with regard to the gaze direction of the user. The ideal vision case is when the gaze direction is perpendicular (normal) to the floating menu surface. The worse vision case is when the gaze is parallel (tangent) to the surface, making it impossible to see the interface elements on the menu surface. For all cases where orientation was predetermined (non controllable by the user), we made so that the menu is perpendicular to the user's gaze (optimal condition). With the reference frame 4 using the left hand, we wanted to let the subject free to do the manipulations while assuming that he would naturally take the most adequate posture with time, according to his training accumulated from his previous trials. Since the left hand case is distinct from the other cases in that regard, it may be a cause of the weaker performances observed with the reference frame 4.

The position of the menus itself was not constant from a case to the other. For example, in the case where the menu is attached to the head movements of the user, the menu is localized to height of the eyes. The user must therefore manipulate the interfaces with the arm raised in front of him. In the cases where the menu was attached to the user's body, the position was lower, at 30 centimetres (1 foot) below the head. In this case, the manipulation was done with the arm placed lower, which causes less muscular fatigue. For the case using the left hand, we observed important variations of the positions of use from

a subject to the other. Some placed the menu at eye height whereas others placed it at waistline height, or halfway between the two.

The distance between the eye and the menu surface is another factor related to the position that was not constant from a case to the other during our experiences. As we have seen it previously, this distance affects the acuteness with which the user will be able to perceive an object in the scene. It was especially notable with the text displayed on the menus. Some subjects indeed made the remark that they had difficulties to read the text on the menus in some circumstances. Reading text inside a helmet remains a critical element limiting the use of such devices in the VR domain. According to our observations, an angular resolution of 1.5 minute of arc per pixel or less is required⁴ to allow reading text with an immersive helmet, or any other display device.

Visual acuity concerns also apply to the buttons and cursor dimension. As a minimal measure, we made them to be larger than the letters used for the texts considering that the text size we used was at the limit of readability inside the helmet used. A more in-depth study about the relation between the visual acuity and the dimensions of the VR interface elements would be interesting to do. It would allow us to get precise metrics and to put in equations the link that exists between these two parameters. The utility of such equations would allow to systematically calculate the optimal dimension of the interface elements according to the display device used.

There is therefore a potential variability factor in our results regarding these positioning and orientation considerations of the menus, from a case to the other. This being said, the position variations were small and do not invalidate our measures. In the cases where the menu had a stationary position for example, the orientation of the menu surface was always so that the gaze was perpendicular to the menu surface and therefore following optimal conditions.

D. Precision vs. user fatigue

According to our results, the user's fatigue affects the precision of manipulation more than the time of manipulation. The precision should therefore receive a particular attention in the design of HMI aimed at VR systems. Our findings would have probably been different if we had not imposed a minimal precision to control the trial ending condition. In this case, we would probably have seen a more direct and proportional interrelationship between the precision and the manipulation time that would have both been more extensively affected. By imposing a minimal cursor releasing precision, we made such that the measures of time made have been controlled in part by the precision. That is to say that the positioning error was not free of all constraint. It was forced to a peak value not representing the true positioning error that one would have measured

⁴ lower minute of arc values represent better perceived resolution (i.e., the visual acuity)

in the absence of this constraint. The time was indirectly affected by this position constraint, as when the cursor was released with too much positioning errors, the chronometer continued to increment the time of the trial and the subject had to seize the cursor again, move it closer to the target and to release it.

Regarding the learning curve, we observe that the subjects are more precise than fast at the beginning of a block of trials. Thereafter, we observe an acceleration with regard to the manipulation times that shows learning and a raise in trust that allows the subject to be faster. On the other end, this diminution of manipulation times comes with a notable precision reduction (increase of the positioning error) starting to manifest halfway through the block of trials that denotes an inversely proportional link between the manipulation speed and the precision. It is in concordance with the Fitts law, well known in the 2D HMI domain. The Fitts law shows that an inversely proportional link exists between the travel time to reach a target and the target dimension. In the present case, we consider the positioning error as being the target dimension. The travel time is the cursor manipulation time, from its initial position toward the final target indicated by the arrow. Therefore in fact, the more the cursor is precisely released and well aligned with the arrow (smaller target), the more the travel time has the tendency to increase.

E. Considerations relative to the subject's gender

Regarding the subject's gender, we observe that women are more meticulous and applied than men. It allows them to be more precise in general. In counterpart, women are slower and have more variable results. A directly proportional link exists between the precision and the variability of the precision measures. The more the subject tries to be precise, the more he manipulates for a longer time and so develops more fatigue. As trials go by, this increased fatigue induces involuntary movements that will have for consequence to produce a bigger variability of the cursor end positions. The more mechanical movements observed with men favour a better reproducibility of the results, this in spite of a lower concentration level that results in a lower precision.

F. User preferences vs. Analysis results

One of the most surprising finding was that the subjective appreciation of the subjects regarding the five reference frames doesn't agree with the quantitative results. Several subjects preferred the reference frame 4 that has been found to be the worse based on performance analysis. It demonstrates that the perception of a man-machine interface from the human's perspective can depend on other factors than speed and precision. The subjective questions asked to the subjects evaluated their general satisfaction rather than their perception of the performances at the time of the trials. A high degree of satisfaction from the user's point of view doesn't necessarily imply better performances, and vice versa.

VIII. CONCLUSION

This experience allowed us to get quantitative and qualitative data on the impact of the VR HMI reference frame on the user's performances and satisfaction. We found that a coupling between the user reference frame and the one of the floating menus is desirable. The number of degrees of freedom common between the floating menu reference frame and the one of the user should be as high as possible. A floating menu that follows the user's body movements will therefore be preferable to a floating menu fixed in relation to the scene.

This study showed that the frames of reference used to display 3D floating menu VRUIs have a significant impact on user performance. Measured user performance gains for time and positioning errors were in the range of 12 to 19%. The current results tend to show that the best reference frames for VRUIs are the ones which have the closest links with the user's body reference frame, namely, frame 3 and 5. This can be attributed to the fact that using such a frame limits the work required by the brain to adapt user's movements for many different active frames of reference.

Moreover, our observations allowed us to note that an important difference exists between the subject's satisfaction level and the experimental measures regarding the HMI manipulation human performances of VR systems. It is therefore important to keep in mind that an important part of the VR HMI design should include a validation of subjective level of satisfaction of the users as much as by quantitative measures on the manipulations times and the position errors.

This clearly shows us that there remains a substantial opportunity to enhance VRUI interaction performance by carefully studying key factors defining the behaviour of these user interfaces. By combining many such enhancements, it can be expected to achieve substantial performance gains.

Our hope is that this work will stimulate further research in that direction. We feel that by studying the user performance with VRUI, we will be able to greatly enhance the usability of such interfaces for VR systems and doing so, actually greatly enhance the usefulness of VR systems in general.

IX. REFERENCES

- [1] Jacoby, R. and Ellis, S., Using Virtual Menus in a Virtual Environment *Proceedings of SPIE: Visual Data Interpretation*, vol. pp. 39-48, 1992.
- [2] Wloka, M. M. and Greenfield, E., "The virtual tricorder: a uniform interface for virtual reality," *Proceedings of UIST '95: 8th ACM Symposium on User Interface and Software Technology*, 14-17 Nov. 1995, pp. 39-40, 1995.
- [3] Mine, M., UNC Chapel Hill Computer Science, Technical Report TR96-029, 1996.
- [4] Stoakley, R., Conway, M. J., and Pausch, R., "Virtual reality on a WIM: interactive worlds in miniature," *Proceedings of the Conference on Human Factors in Computing Systems. Part 1 (of 2)*, May 7-11 1995, pp. 265-272, 1995.

- [5] Coninx, K., Van Reeth, F., and Flerackers, E., 2D human-computer interaction techniques in immersive virtual environments *Computer Networks and ISDN Systems: Proceedings of the 1996 5th International Conference on Computational Graphics and Visualization Techniques, COMPUGRAPHICS, Dec 16-18 1996*, vol. 29, no. 14, pp. 1685-1693, 1997. 0169-7552. LUC - Expertise Cent for Digital Media, Diepenbeek, Belg.
- [6] Kim, N., Kim, G. J., Park, C.-M., Lee, I., and Lim, S. H., Multimodal menu presentation and selection in immersive virtual environments *Proceedings - Virtual Reality Annual International Symposium: IEEE Virtual Reality 2000, Mar 18-Mar 22 2000*, vol. pp. 281 2000. 0-7695-0478-7. Pohang Univ of Science and Technology, South Korea.
- [7] Lindeman, R. W., Sibert, J. L., and Templeman, J. N., "The effect of 3D widget representation and simulated surface constraints on interaction in virtual environments," *IEEE Virtual Reality 2001, Mar 13-17 2001*, pp. 141-148, 2001.
- [8] Angus, I. G. and Sowizral, H. A., "Embedding the 2D interaction metaphor in a real 3D virtual environment," *Stereoscopic Displays and Virtual Reality Systems II, Feb 7-8 95*, pp. 282-293, 1995.
- [9] Lemoine, P., Vexo, F., and Thalmann, D., "Interaction Techniques: 3D Menus-based Paradigm," *AVIR 2003*, 2003.
- [10] Larimer, D. and Bowman, D., "VEWL: A Framework for Building a Windowing Interface in a Virtual Environment.," *Proceedings of INTERACT: IFIP TC13 International Conference on Human-Computer Interaction*, pp. 809-812, 2003.
- [11] Kohli, L. and Whitton, M., "The haptic hand: Providing user interface feedback with the non-dominant hand in virtual environments," *Proceedings - Graphics Interface 2005, May 9-11 2005*, pp. 1-8, 2005.
- [12] Dachsel, R. and Hübner, A., "A Survey and Taxonomy of 3D Menu Techniques," *Proceedings of the 12th Eurographics Symposium on Virtual Environments (EGVE'06)*, Lisbon (Portugal), 2006.
- [13] Bowman, D. A. and Hodges, L. F., Visualization and Usability Center, Technical Report GIT-GVU-95-26, 1995.
- [14] Donald A. Norman, *The Design of Everyday Things* Anonymous1988. Doubleday. New York. 0465067107.

Marc Bernatchez completed a bachelor and a master's in electrical engineering at Laval University in Quebec, Canada, in 1997. He is now completing a Ph.D. in electrical engineering at Ecole Polytechnique de Montreal, in Montreal, Canada. His main field of interest is virtual reality with a particular emphasis on VR user interfaces.

He worked as a Software Engineer in the VizSim (visual simulation) domain for five years while at CAE Inc., a flight simulation company that engineers commercial and military full flight simulators. He participated on the development of new visual effects in commercial IGs (image generators) such as the MAXFOG™ layered fog engine. He also worked on military projects such as ESVS (Enhanced and Synthetic Vision System) jointly developed by CAE and the NRC-CNRC (National Research Council of Canada Institute for Aerospace Research), that allows rescue helicopter pilots to fly in near zero visibility climatic conditions. His current research interests are virtual reality user interface development and optimization.

Jean-Marc Robert holds a bachelor and Master's degrees in psychology from the University of Montreal in Quebec (1976) and a Ph.D. in human factors from the University Paris V (La Sorbonne) in France (1982). He also did an internship at Xerox-PARC (1983) and 2-year post-doctoral studies at NASA-Ames Research Center (1983-1985) in U.S.A. under a grant from the National Science Foundation. His main field of study is human-computer interaction, collaborative work, and the use of technology in training.

He is full professor at Ecole Polytechnique de Montréal in Quebec where he teaches cognitive ergonomics and engineering and human-computer interaction. He is also the founder and the director of LIHM, a research laboratory on Human-Machine Interaction at École Polytechnique. .