

Did “Minority Report” Get it Wrong? Superiority of the Mouse over 3D Input Devices in a 3D Placement Task

François Bérard¹, Jessica Ip², Mitchel Benovoy²,
Dalia El-Shimy², Jeffrey R. Blum², Jeremy R. Cooperstock²

¹ University of Grenoble,
Laboratoire d'Informatique de Grenoble,
BP 53 - 38041 Grenoble cedex 9, France
Francois.Berard@imag.fr

² McGill University,
Centre for Intelligent Machines,
Montréal, H3A 2A7, Canada
{jessicaip|benovoy|dalia|jeff|jer}@cim.mcgill.ca

Abstract. Numerous devices have been invented with three or more degrees of freedom (DoF) to compensate for the assumed limitations of the 2 DoF mouse in the execution of 3D tasks. Nevertheless, the mouse remains the dominant input device in desktop 3D applications, which leads us to pose the following question: is the dominance of the mouse due simply to its widespread availability and long-term user habituation, or is the mouse, in fact, more suitable than dedicated 3D input devices to an important subset of 3D tasks? In the two studies reported in this paper, we measured performance efficiency of a group of subjects in accomplishing a 3D placement task and also observed physiological indicators through biosignal measurements. Subjects used both a standard 2D mouse and three other 3 DoF input devices. Much to our surprise, the standard 2D mouse outperformed the 3D input devices in both studies.

Keywords: 3D placement, input device, stress.

1 Introduction

The growing availability of low-cost graphics rendering hardware and associated 3D applications targeting non-specialists (e.g., Google Earth, Sketchup) has focused increased attention on the problem of improved interaction paradigms for 3D tasks, for example, positioning an object in a volume. Such tasks are inherently more difficult to perform than 2D tasks because of the additional degree of freedom in which users must operate. However, we live and interact with objects easily and efficiently in the 3D physical world. Thus, it seems natural that computer systems should be able to leverage these physical-world skills for interacting with 3D virtual worlds. This intuition is expressed in motion pictures such as “Minority Report”, portraying compelling scenarios of possible 3D interactions. To make this transfer of skill possible, new devices have been developed that

better match the high DoF of our hands than the mouse does. Indeed, many prototypes have been proposed in the literature [1-3], some of which have evolved into commercial products (e.g., the *SpaceNavigator*^{TM1}). Nevertheless, despite the wealth of research and development into 3D input devices, the mouse remains the dominant input device for performing 3D tasks, at least in the case of desktop applications such as Computer Aided Design (CAD) or 3D modelling.

This paper presents an effort to understand the preference for the mouse. Specifically, we attempt to answer the fundamental question: does the mouse persist as the dominant input device for 3D tasks because of factors beyond performance, for example, cost or ingrained preferences, or is it the intrinsic qualities of the mouse that make it better suited for 3D interaction than high DoF devices? While the latter possibility appears counterintuitive, it is worthwhile to investigate because it could have significant impact on 3D interaction research.

This question, however, must first be expanded. Interaction is not defined by an input device alone, but by the combination of a device and an interaction technique. In the example of 3D object rotation, the mouse is typically used with the virtual sphere technique [4] while a free-space device is used with a direct mapping (either absolute or relative) [3]. Thus, each device must be matched with its most suitable interaction technique in making performance comparisons, rather than choosing a single interaction technique for all devices. Moreover, “3D interaction” encompasses a large spectrum of tasks, including simple object picking to more complex 6 DoF docking. Different combinations of device and interaction technique may be better suited for some of these tasks, but not for others. However, regardless of the diversity of tasks and possible choices of device and interaction technique, users have primarily settled on the mouse, so the fundamental question raised above remains valid.

In the remainder of this paper, we first review related literature and introduce our methodology for investigating this question. In particular, we justify our choices of benchmark task, set of devices and interaction techniques. We then report on two experiments comparing user performance on a 3D object placement task with a mouse and three other higher DoF devices. Finally, we offer some conclusions and suggest directions for possible further research.

2 Related work

Several proposals for a suitable device for object placement in 3D can be found in the literature. These include the *Bat* [3], a free-space, 6 DoF device held in the hand; the *Roller mouse* [2]: a regular mouse augmented with wheels on top of the case that add one DoF for the control of cursor depth; the *Rockin'mouse* [5], another augmentation of a regular mouse that tilts around its base to provide one or two additional DoFs; and the *GlobeFish* and *GlobeMouse* [1], which decouple translation and rotation at the device level: translation is isometric, or elastic rate-control, and rotation is isotonic using a trackball. The performance of most of these devices has been the subject of formal user experiments. Unfortunately, fair comparisons between them are rare as the experiments

¹ <http://www.3dconnexion.com/3dmouse/spacenavigator.php>

tend to differ from each other in the performed task, and most experiments neglected to include the mouse in the comparisons.

Previous studies have shown that in some situations, the mouse is less efficient than higher DoF devices. McMahan et al. [6] found that the mouse was less efficient than 6 DoF devices on a 6 DoF task (docking). Hinckley et al. [7] show that a free-space device is more efficient than mouse-based interaction for object rotation, a 3 DoF task. We chose placement rather than docking and rotation because we consider it a more fundamental and frequent task, as described in further detail in Section 3.1 below.

Balakrishnan et al. observed that users typically interact both in 2D and 3D, and 3 DoF (or greater) devices are usually poorly suited for 2D interaction. This motivated their development of the *Rockin'mouse*, a 3D device modelled after a regular mouse, but which provides an additional DoF from its tilting base. The *Rockin'mouse* was shown to be superior to the mouse [5] for placement tasks, although this result relied on the use of a non-standard interaction technique that was arguably biased toward a 3 DoF input device. We go one step further in examining the possibility that the regular 2D mouse could actually be *more efficient* for 3D object placement than higher DoF devices, regardless of the 2D context in which 3D interaction takes place. We evaluate this possibility by testing the mouse in conjunction with its best-suited interaction technique: a conventional 4-views display environment.

More recently, Teather et al. [8] tried to discover what properties of the mouse make it a *good* device for 3D placement. They found that the mouse outperformed a 3D tracker (Intersense IS900) when the tracker was used as a “mouse emulator” (i.e., constrained to a 2D surface). They hypothesized that the higher resolution of the mouse could be the key factor in its superiority. However, the tracker was used in *absolute* mode (like a tablet) while the mouse was used in its natural *relative* mode; this may have contributed to the observed difference in performance. They also found that both the mouse and the “mouse emulation” tracker outperformed the tracker when used as a free-space device, although the interaction technique was based on sliding one object over top of other objects in the scene, thus reducing the task from 3 DoF to 2 DoF. We are interested in evaluating the performance of the mouse on a true 3 DoF task without constraining the necessary operations to a 2D surface.

In the following section, we describe our methodology in trying to compare, as objectively as possible, the performance of the mouse and higher DoF devices in a 3 DoF object placement task.

3 Methodology

Our goal was to establish whether the dominance of the mouse is due to its intrinsic qualities for 3D interaction, or to *other factors*: its broad availability, the mixing of 2D and 3D tasks in user's typical work, or the long user habituation to the mouse. We conducted a set of studies comparing user performance in a 3D task while using the mouse and a set of credible, representative examples of 3 DoF input devices. We designed these studies so that pure 3D performance was isolated from the *other factors*: device availability was obviously not an issue in a controlled experiment, the experiments only involved a pure 3D task, and the experiments were designed to last sufficiently long so that the learning curve for each device could be inferred, providing an indication of device performance

after user training. We also aimed at a direct measurement of device effect on stress through the recording and analysis of the participants' biosignals during the experiment. We now explain our rationale for selecting the benchmark task, the tested devices, the interaction techniques, and the choice of recorded biosignals.

3.1 The Task

One of the most frequent tasks in 2D graphical interaction is the selection of objects through spatial acquisition. This suggests, similarly, that target selection in 3D scenes would be a good candidate task. However, as the 3D scene is always perceived through 2D projections (on the screen or on the retina), target selection in 3D can be simplified to a 2D task. In the physical world, this is done by closing one eye and pointing at an object. In the digital world, the equivalent is 2D selection of the *projection* of the object on the screen, otherwise known as “picking”. There is little incentive, then, to use a high DoF device when the mouse is available and has been highly optimized for this task, especially when combined with smart selection techniques such as the bubble cursor [9], or semantic pointing [10].

Another frequent task is the placement of objects at arbitrary locations. This may involve changing the position of some object in a 3D modeling application, or positioning a marker within volumetric medical data. In either case, the system is unable to provide the assistance it can with picking tasks, because it has no clue as to where the user wants to place the object. The user is thus required to specify all three coordinates in some manner: in other words, this is a true 3D task. Indeed, there is presently no way to perform this task directly with a regular mouse. Instead, the standard solution is to decompose the task into two 2D placements, e.g., one in the x - y plane, and a second in the x - z plane. This suggests that 3 DoF devices should prove superior in performance by supporting task completion in a single action.

3D placement can be reduced to a 2 DoF task by the clever use of constraints: objects are, most of the time, placed next to each other. They are not left floating in the air. Hence, placement can be performed by sliding the moving object on top of the other (static) objects in the scene [8,11]: a 2 DoF task. While this kind of clever interaction technique illustrates the possibilities of performing 3D tasks with a 2D mouse, we want to compare the efficiency of the mouse with higher DoF devices in a true 3D task. For this reason, we chose the unconstrained positioning task in our experiment.

Other candidate tasks involving three or more DoF include object rotation, docking (a 6 DoF object placement including both a translation and a rotation), and navigation, which is equivalent to docking the camera that is viewing the scene. Our rationale for selecting placement as the benchmark task is as follows: if a number of 3D input operations are required by the application, we believe that users will choose the device that is best suited for the most fundamental and frequent of these (i.e., placement). Furthermore, previous studies [12] have shown that coordinating 6 DoF is too difficult to perform in a single action, and thus, 6 DoF tasks such as docking are actually performed by a sequence of 3 DoF translations (placements) and rotations even on 6 DoF devices.

3.2 The Devices

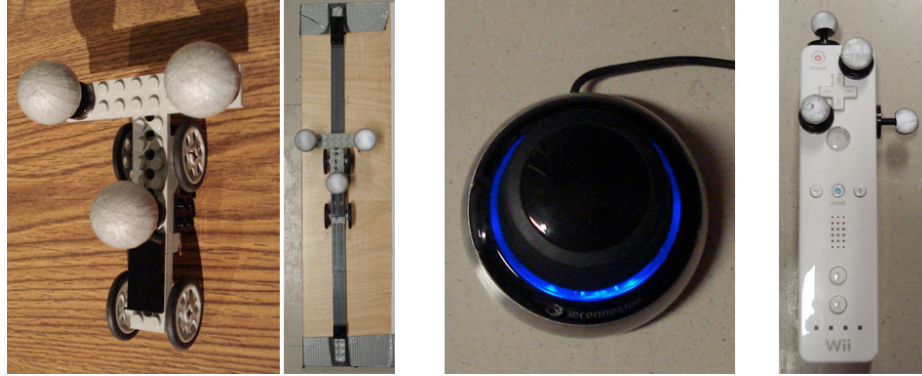


Fig. 1. The devices (mouse not shown). From left to right: the *DepthSlider* alone and on its rail, the *SpaceNavigator*TM, and the free-space device *Wii Remote*TM tracked by a *Vicon*TM.

Our objective was to compare the mouse against devices that are considered state of the art for 3D input. In addition to the mouse, we ran the experiment with a free-space device, a *SpaceNavigator*TM, and the *DepthSlider*, a mouse augmentation we designed to exploit bi-manual input. These are similar in design to systems described in previous literature, as summarized in the table below, and illustrated in Figure 1.

Table 1. Tested devices and related systems

Tested device	Properties	Related system
Mouse	2 DoF isotonic	
<i>DepthSlider</i> (mouse+slider)	2+1 DoF, isotonic	Roller Mouse and Rockin'mouse
<i>SpaceNavigator</i> TM	6 DoF elastic	GlobeFish and GlobeMouse
free-space	6 DoF isotonic	Bat

The DepthSlider The *DepthSlider* is a physical slider, whose base is attached to the desk on the side of the keyboard opposite the mouse, and positioned so that its principal axis of movement is toward or away from the screen. It is inspired by the *Roller Mouse* [2] and the *Rockin'mouse* [5] in the sense that the objective is to augment the mouse to better support 3D interaction, rather than replace it. It differs from the earlier devices in that the mouse is not physically modified, and thus retains all the properties that make it suitable for 2D interaction. The *DepthSlider* is therefore used by the non-dominant hand to control depth, while the mouse is used with the dominant hand in the standard way (control of x and y values). The slider operates in *relative* mode, the same way as the mouse. The depth of an object is only controlled by the slider when the mouse “holds” it, that is, with the pointer on the object and the mouse button depressed. In consequence, the mouse controls the clutching and un-clutching of the slider. Un-clutching is necessary when the slider reaches either of its physical limits: the object is then released so that the slider can be moved away from the limit. Clutching again allows the user to resume moving the object.

The SpaceNavigator™ The *SpaceNavigator* (or simply, the *Navigator*) is a widely used, commercial device with 6 DoF. It provides elastic rate-control, which is considered optimal for coordination and minimizing fatigue [13].

The *Navigator* is intended for desktop use in the non-dominant hand so that the user does not have to switch hands to operate the mouse. In our experiments, the *Navigator* is used in this configuration, while the dominant hand rests on a mouse for target validation (pressing the right mouse button). We only use the 3 DoF translational control of the *Navigator*, activated by applying lateral forces to the device. Reducing the DoF to translation only is in fact common and supported explicitly by the device driver, as this helps avoid unintended translations resulting from rotation, and vice versa.

The free-space device We designed a free-space device, inspired by the *Bat* [3], using the Nintendo *Wii Remote™*, with the addition of infrared reflectors to track translation and rotation with a *Vicon™* motion capture system. The remote was selected mainly as an ergonomic hand-held device, but its buttons were also used for clutching (“B” trigger button below the remote) and to indicate task completion (“A” button on top).

3.3 Interaction Techniques

We expected performance to vary significantly depending on the choice of interaction technique. Since our study concerned the dominance of the mouse in 3D interaction, our approach was to study first whether or not the most *common interaction technique* used with the mouse was the key factor. We did this by conducting an initial experiment where all devices used the same (common 4-view) interaction technique. In a second experiment, we tried to optimize the interaction technique used with each device.

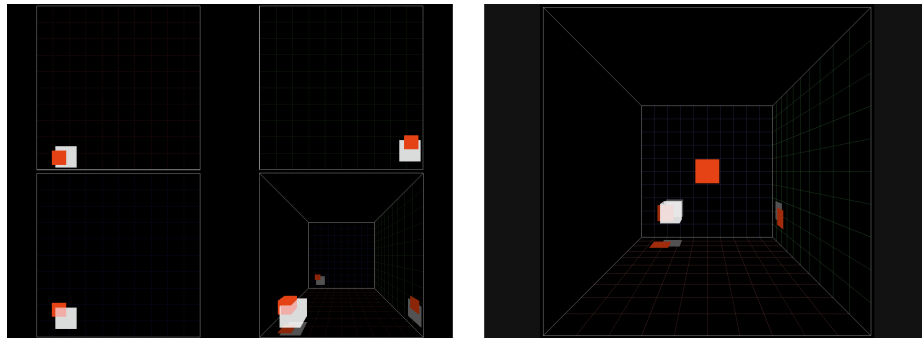


Fig. 2. Graphical display. Left: 4-view mode (top, side, front, and perspective view of the scene). Right: fullscreen perspective mode. The translucent white cube is the controlled object; the opaque red cube is the target location.

Subjects were asked to move the *object* (the white cube in Figure 2) so as to enclose the *target*, represented by the red cube, and confirm their placement by pressing a button (the right mouse button for all devices except for the free-space device, which used the *Wii Remote™* “A” button). The target had to be completely surrounded by the object for the placement to succeed. Otherwise, when the button was pressed, graphical feedback

indicative of the error was provided by changing the color of the white cube to red for one second. Object and target were both represented as cubes and the object was made semi-transparent to provide the benefits of the Silk Cursor [14]. Both cubes cast “shadows” on the sides of the scene as additional location cues.

First experiment The scene was rendered in four different views as illustrated in Figure 2, left. Three of the views were orthographic projections allowing interaction in two dimensions of the scene (front: x - y , top: x - z , and side: y - z). The fourth view displayed a more natural perspective rendering of the scene.

With the mouse and the *DepthSlider* (which both involve the mouse for placement), the object must first be acquired before being moved. Object acquisition is done in the classical way: the mouse *pointer* is moved on over the object, the left mouse button is depressed, and the object is then coupled with the device such that its position is controlled until the left mouse button is released. In the mouse condition, it is necessary to execute a minimum of two sequences of object acquisition and placement in two different views because each view only allows the control of 2 DoF in the 3 DoF task. In the *DepthSlider* condition, only one object acquisition is necessary because the mouse controls the x/y position of the object while the slider controls the z position. In contrast with the mouse and the *DepthSlider*, the *Navigator*TM and the free-space device are always coupled with the object. No acquisition is necessary as the object moves as soon as the devices are controlled by the user. The rationale is that these devices are specialized for 3D operations; hence, they are meant to be used in conjunction with a mouse for general 2D operations such as object selection (picking).

The free-space device is used in *absolute* mode: the mapping of the physical location of the device to the virtual location of the object is fixed. The control/display (C/D) ratio is empirically chosen once for all experiments, and the origin of the mapping is re-defined before each experimental session so that participants can comfortably reach every location of the virtual scene.

Second experiment After running the first experiment and seeing the results, we realized that the four views were not a good match for the high DoF devices. The orthographic views encourage the decomposition of motion into 2D segments, but the 3D devices should be more efficient when all dimensions of the task are solved in a coordinated manner. We thus designed the second experiment where all devices but the mouse are used with a more natural, full screen, perspective view of the scene (Figure 2, right). In addition, the view is rendered in stereo in order to improve depth perception.

Object acquisition was the same as before, but we introduce a new *relative* mode of operation for the free-space device. When used in relative mode, the object is only coupled to the object when the clutching button of the *Wii Remote*TM is depressed (the “B” trigger situated below the remote), as with the *Bat* [3]. This permits a dynamic C/D ratio: we set a strong gain during fast motion to allow quick positioning during the early phase of motion, and the gain is reduced to a small value during slow motion to facilitate accurate positioning. A dynamic C/D ratio is the main motivation for the introduction of the relative mode, as we observe in absolute mode that it is not possible to provide both fast initial positioning and accurate final positioning.

3.4 Biosignals as a measure of stress

By recording the biosignals of the participants while performing the task, we obtain a continuous measure of the level of stress induced over time by the input device being used. We believe that the use of biosignals to measure stress presents a novel and practical way to help gauge the usability of devices. We accomplished our stress assessment by using three measures of autonomous nervous system activation, namely galvanic skin response (GSR), blood volume pulse amplitude (BVP) and heart rate (HR).

Galvanic skin response (GSR) measures the electrical conductance of the skin. The signal can be decomposed into skin conductance responses (SCR), related to short events, and the skin conductance level (SCL), related to the underlying basal arousal activity. The GSR is often the primary psychophysiological measure used when gauging emotional activation as it responds very quickly (1-3 seconds after onset of stimulus) [15] and correlates highly with stress or anxiety. The blood volume pulse (BVP) signal is an indicator of blood flow using a photoplethysmography [15]. The amplitude of the blood volume pulses tends to decrease following sympathetic arousal [16,17]. The heart rate (HR) is computed from the raw BVP waveform by finding consecutive local maxima. An increase in sympathetic activity will increase the heart rate. In summary, higher stress is detected with lower BVP values and higher GSR, SCR, and HR.

4 First Experiment

4.1 Goal and hypothesis

The goal of the initial experiment was to determine which device is the most efficient and causes the least fatigue during 3D placement. Our hypotheses were as follows:

- H1a: some of the present authors anticipated that direct specification of 3D position by hand movement in free space would outperform all other methods. The apparent naturalness of placing physical objects in a similar manner suggested that such interfaces represented an obvious improvement over devices constrained to operations in a 2D plane such as a mouse.
- H1b: other authors expected that the *DepthSlider* would be the most efficient device because they assumed that manipulating a stable device on the table surface would be key to task performance, outweighing the benefits of the natural mapping of a free-space device. They further expected that two-handed interaction could be quickly mastered, as shown in previous studies, and that the additional DoF offered by the slider would improve performance compared to the mouse, alone.
- H2: we expected to measure signs of fatigue from use of the free-space device because it does not allow the hand and the arm to rest on the table.

4.2 Task and Apparatus

Participants were asked to place a translucent white cube so as to enclose a red target cube, as described in Section 3.3. Subjects sat at a desk, facing an LCD monitor. In the free-

space condition, the desk was removed, leaving the dominant hand able to move freely. A Vicon motion capture system tracked the free-space device in 6 DoF and the 1 DoF of the *DepthSlider*, using infrared-reflective markers attached to each device.

Graphical feedback (Figure 2, left) was displayed on an LCD screen (1280x1024 pixels resolution, 19" diagonal with a 4:3 aspect ratio, 60 Hz refresh rate), using custom-built software on top of the OpenSceneGraph library (<http://www.openscenegraph.org/>).

Physiological data was collected with Thought Technology's biofeedback system (<http://www.thoughttechnology.com>). All signal were measured by sensors strapped to fingertips. For consistency, the sensors were placed on each participant by the same experimenter. As a normalization step, the average of the one minute silent baseline data preceding each trial was subtracted from the “active” data of the participant. The data was then divided by the range of the participant's active data (i.e., $max - min$) to account for inherent physiological differences between participants.

4.3 Design and procedure

In order to study the effect of device usage duration on performance and induced stress, trials were grouped into six consecutive *trial blocks*. We used within-subject design with independent variables *input device* (mouse M, Navigator N, DepthSlider D, and free-space Wii remote W), and *trial block* (1st, 2nd, ..., 6th). Each trial block consisted of 3 *runs*. Each run consisted of 9 targets combining 3 Fitt's indices of difficulty ($ID = 1, 2, 4$) with 3 target widths (7, 14, and 22 pixels). The orientation towards the target was randomized, as was the target's presentation order within blocks, but the same random sequence was used for all conditions and all participants. The order of device presentation was balanced between all participants.

Ten participants (5 male, 5 female) took part in the experiment. All were members of our laboratory, right-handed, and none had experience with the *Navigator*, the free-space device or the *DepthSlider*. The experimental design was as follows: 10 participants x 4 devices (mouse, Navigator, DepthSlider, free-space) x 6 blocks of trials x 27 targets = 6480 *placements* in total. Participants took part in four sessions (one for each device). Sessions lasted approximately half an hour each and were separated by at least four hours to ensure sufficient recovery time from any fatigue induced by the previous session. At the beginning of a session, the task and the mode of operation of the tested device were explained verbally. Participants were told that they should perform the task as fast as possible while trying not to make placement errors. They familiarized with the device on the first 36 trials (4 runs) which were used for training. The actual trial blocks were then executed immediately after the training with no rest between the training and the first block, nor between the 6 blocks. The goal was to create a task that was sufficiently demanding for fatigue to be measured. The dependent variables measured were placement time, placement error, mean blood volume pulse amplitude, mean heart rate, mean galvanic skin response, and the number of SCRs.

4.4 Results

Figure 3 (top-left) represents the mean of total session time and mean of error rates across all participants for each device. An ANOVA on the data shows a strong main effect of

device on placement time ($F_{3,9}=10.7, p < 0.001$) and error rate ($F_{3,9}=72.9, p < 0.001$). Post-hoc analysis shows that both the mouse and *DepthSlider* are significantly more efficient ($p < 0.01$) than the *Navigator* (30% faster) and the free-space device (40% faster), but the difference was not significant between the mouse and the *DepthSlider*, nor between the *Navigator* and the free-space device. Participants made significantly more errors in the free-space condition than in all other conditions ($p < 0.001$). The only other significant difference in error rate was between the mouse and *Navigator* conditions ($p < 0.05$).

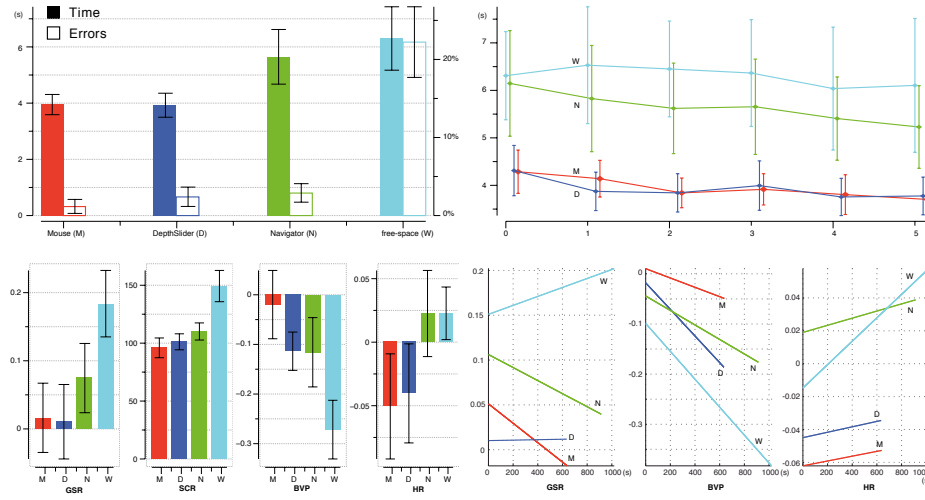


Fig. 3. Results of first experiment. All figures with 95% confidence intervals. *Top-left:* Means of average placement time (left axis, in seconds) and error rate (right axis, in %) per device across all participants. *Top-right:* Means of average placement time (s) per condition and trial block across all participants. *Bottom-left:* means of normalized conductance (GSR), #skin conductance responses (SCR), normalized amplitude (BVP), and normalized heart rate (HR) per device across all participants. *Bottom-right:* Linear regressions of the mean biosignals waveform across all participants over task completion time (in s) per device (same scales as bottom-left).

Figure 3 (top-right) shows the evolution of trial block completion time across the six blocks. There was a significant effect of block number on completion time in the mouse condition ($F_{5,9}=14.1, p < 0.001$), *DepthSlider* condition ($F_{5,9}=7.79, p < 0.001$) and the *Navigator* condition ($F_{5,9}=6.89, p < 0.001$), but not the free-space condition ($F_{5,9}=0.944$). There was no significant effect of block number on number of errors.

The results of the mean physiological data for each input modality across participants is shown in Figure 3 (bottom-left). An ANOVA yielded a significant difference for the GSR ($F_{3,36}=2.89, p < 0.05$), #SCRs ($F_{3,36}=7.41, p < 0.01$), BVP ($F_{3,36}=3.9, p < 0.05$), HR ($F_{3,36}=3.72, p < 0.05$). Figure 3 (bottom-right) shows the mean GSR, BVP, and HR data across all the participants for each input device type as it varies over the time of the experimental session. The data were first re-sampled to the average time-of-completion for the given device, then fit to a line with a robust linear regression ($r > 0.59, p < 0.05$ across all regressions) in order to reveal the underlying timewise trend. The most noticeable result is that the free-space device led to the greatest stress as it induced the highest mean GSR and #SCRs, and the lowest mean BVP. It also produced higher HR than the mouse and the *DepthSlider*, but for this measure, there was no significant difference with the

Navigator. For the GSR, the mouse and *Navigator* have a downward trend, indicating that the subjects' stress may be decreasing as time proceeds while using these devices. Conversely, the free-space device seems to induce progressively higher GSR. The higher rates of decline of BVP and increase of HR indicate a more powerful effect of the free-space device on stress.

5 Second experiment

5.1 Objectives and variations from the first experiment

The main goal of the second experiment was to test whether the previously observed superiority of the mouse holds when high DoF devices were used with a graphical display that better promotes coordinated input than the 4-views layout, in particular, a stereoscopic display in which depth can be observed directly. A secondary goal was to evaluate if the free-space device would perform better in *relative* than *absolute* mode. The two experiments were similar in design apart from the following:

- *Graphical display*. Graphical feedback was projected on a screen approximately 1.5 m from the participant's chair. In the mouse condition, a single projector (1024x768 resolution) provided the same feedback as in the first experiment, albeit at reduced resolution. In the other conditions, an additional 1024x768 resolution projector was used, in combination with polarized filters and glasses, in order to create a stereo rendering of a single, fullscreen, perspective view of the scene.
- *Design*. As the first experiment indicated quick learning of the new devices, mostly during training, we did not feel it necessary to reproduce the long experimental sessions, hence only three blocks of trials were used. A relative mode was added for the free-space device, as well as a higher Fitts' Index of Difficulty ($ID = 6$). We situated the targets along the eight diagonals from the center (initial position) to the workspace vertices to control the orientation to target. Ten right-handed participants (8 male, 2 female) took part in the experiment; all were members of our laboratory. The experimental design consisted of 10 participants x 5 devices (mouse M, DepthSlider (Stereo) DS, Navigator (Stereo) NS, free-space absolute (stereo) WS, free-space Relative (Stereo) WRS) x 3 blocks of trials x 4 indices of difficulty (1, 2, 4 and 6) x 8 diagonal orientations = 4800 *placements* in total.

5.2 Results

Figure 4 describes the results of the second experiment. An ANOVA revealed a significant effect of device on task accomplishment time ($F_{4,9}=5.83$, $p < 0.01$) and error rate ($F_{4,9}=17.6$, $p < 0.01$). The *Navigator* was significantly slower than all other devices. Contrary to the first experiment, there was no significant difference in time between the mouse and the free-space device in either absolute or relative mode. However, error rates were significantly higher with all devices compared to the mouse. The best competitor, the free-space device in relative mode, had a 22.6% error rate, more than three times higher than the mouse, at 6.04%.

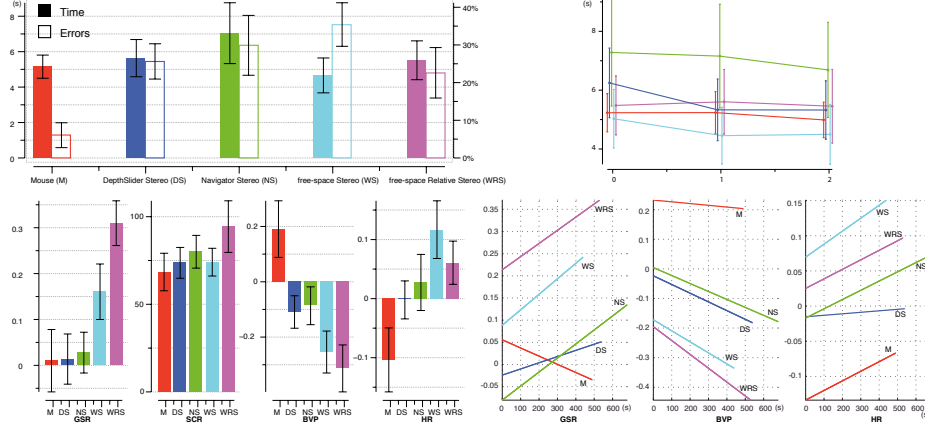


Fig. 4. Results of second experiment. Legend is the same as Figure 3.

An ANOVA on the physiological results revealed a significant effect in the GSR ($F_{4,45}=6.64, p < 0.01$), BVP ($F_{4,45}=4.58, p < 0.01$), and HR ($F_{4,45}=4.01, p < 0.01$), but not in the SCRs ($F_{4,45}=2.82, p > 0.05$). For the GSR, there is again a clear distinction between the tabletop devices and the two modalities for the free-space device. The latter produced significantly higher GSR, indicating higher stress, with the relative mode the highest. The BVP results are similar: the free-space device produced the lowest BVP amplitude, another indicator of stress, whereas this value was significantly higher with the mouse than the two other tabletop devices. The mouse also induced the lowest heart rate. The mean trends shown in Figure 4, bottom-right, ($r > 0.65, p > 0.05$ across all regressions) indicate that the stereo visualization mode may induce higher stress levels over time as all the GSR slopes except the mouse (which was not viewed in stereo) are positive.

6 Discussion

The most natural device (free-space) suffered the worst performance in terms of time (first experiment) and error rate (both experiments). The task was particularly difficult to perform with the free-space device, especially when maximal accuracy was required to position the center of the object in a $4 \times 4 \times 4$ pixel cube. Several factors may explain this difficulty: tracking jitter, parasitic motion induced by clicking the button to terminate the placement, hand instability in free space, and fatigue. In the second experiment, more natural motion trajectories were performed, since these did not need to be decomposed into the 4-view display planes. This improved speed but decreased accuracy. Since the results contradict our hypothesis H1a, we considered whether the task might be biased against the free-space device, or whether this device is not designed for accurate placement. Indeed, more complex 3D tasks, such as docking or deformation, might yield completely different results. However, we believe that the placement task is a fundamental building block of other more complex tasks, and thus the lack of accuracy observed here will likely translate to other tasks. The adequacy of the free-space device appears to depend more on the required accuracy than on the nature of the task.

To determine whether motion tracking was the source of inaccuracy, we recorded the static position of the free-space device on a table for three minutes, then repeated the measurement with five participants attempting to hold the device stationary, without support. Analysis revealed that our Vicon system offered stability of 0.034 mm (equivalent to 740 dpi) whereas the most stable participant maintained the device within an average radius of 0.15 mm over successive time intervals of 0.5 s (the mean for the five participants was 0.48 mm). This indicates that the lack of accuracy is due mainly to the difficulty of holding the free-space device stable in the air. As such, desktop devices (mouse, *DepthSlider* and *Navigator*) benefit from the stabilizing effect of the hand resting on the desk while the fingers execute precise control. This explains the low error rate (< 3%) of desktop devices in the first experiment compared to the free-space device (22%). However, in the second experiment, the *DepthSlider* and *Navigator* desktop devices had error rates over 25% when used with a single perspective view display, whereas the mouse remained at 6% with the 4-views display. While the perspective view was more natural and encouraged coordinated (i.e., faster) motion, it appeared inferior for placement evaluation. We did not consider the level of indirection in the system. In all our experiments, placement is done through *tele-operation*, with the placed object distant from the hand performing the action. Different results may be obtained with the object co-located with the hand (e.g., through a head-mounted display), as this would allow proprioception, but would not help to stabilize the hand.

In summary, we find that both the most natural device and most natural display (i.e., most resembling properties of the physical world) are not the best suited for accurate placement tasks. On the contrary, less natural tools may offer important attributes (e.g., accuracy for the mouse, optimized viewpoints for the 4-views) that are better suited for some tasks. Naturalness may offer strong benefits for immersion and ease of use, but our study indicates that it can also decrease performance when accuracy is required. While this may seem predictable a posteriori, it contradicts the strong intuition that more “natural” interaction is systematically superior.

Surprisingly, the *DepthSlider* was not faster than the mouse, contradicting our hypothesis H1b. While participants did not demonstrate any difficulty with the two-handed interaction, the task did require significant accuracy from the non-dominant hand, which could explain the higher level of stress induced compared with the mouse.

7 Conclusion

We conducted two experiments intended to study the dominance of the mouse in desktop 3D interaction. We compared both user performance and biosignals measured on a fundamental 3D task: object placement in the scene, with the traditional 2D mouse and various 3 DoF devices. Contrary to our expectations, the experiment demonstrated that despite the lack of a third degree of freedom, the mouse was more efficient than the 3 DoF devices for accurate placement. Its accuracy and low level of induced stress appear to compensate for the requirement to decompose 3D tasks into several 2D tasks. The experiments also demonstrated, through biosignals analysis, what had been hypothesized but never shown in the literature: that the free-space devices induce significantly more stress than desktop devices. This indicates that biosignals analysis can be a powerful addition to the set of tools used in usability evaluation.

References

1. Froehlich, B., Hochstrate, J., Skuk, V., Huckauf, A.: The globefish and the globemouse: two new six degree of freedom input devices for graphics applications. In: ACM Conference on Human Factors in Computing Systems (CHI), ACM (2006) 191–199
2. Venolia, D.: Facile 3d direct manipulation. In: ACM Conference on Human Factors in Computing Systems (CHI), ACM (1993) 31–36
3. Ware, C.: Using hand position for virtual object placement. *The Visual Computer* 6(5) (1990) 245–253
4. Chen, M., Mountford, S.J., Sellen, A.: A study in interactive 3-D rotation using 2-D control devices. *ACM SIGGRAPH Computer Graphics* 22(4) (1988) 121–129
5. Balakrishnan, R., Baudel, T., Kurtenbach, G., Fitzmaurice, G.: The rockin'mouse: integral 3d manipulation on a plane. In: ACM Conference on Human Factors in Computing Systems (CHI), ACM (1997) 311–318
6. McMahan, R.P., Gorton, D., Gresock, J., McConnell, W., Bowman, D.A.: Separating the effects of level of immersion and 3D interaction techniques. In: ACM Symposium on Virtual Reality Software and Technology, ACM (2006) 108–111
7. Hinckley, K., Tullio, J., Pausch, R., Proffitt, D., Kassell, N.: Usability analysis of 3d rotation techniques. In: ACM Symposium on User Interface Software and Technology (UIST), ACM (1997) 1–10
8. Teather, R.J., Stuerzlinger, W.: Assessing the effects of orientation and device on (constrained) 3D assessing the effects of orientation and device on (constrained) 3d movement techniques. In: IEEE Symposium on 3D User Interface. (2008) 43–50
9. Grossman, T., Balakrishnan, R.: The bubble cursor: enhancing target acquisition by dynamic resizing of the cursor's activation area. In: ACM Conference on Human Factors in Computing Systems (CHI), ACM (2005) 281–290
10. Blanch, R., Guiard, Y., Beaudouin-Lafon, M.: Semantic pointing: improving target acquisition with control-display ratio adaptation. In: ACM Conference on Human Factors in Computing Systems (CHI), ACM (2004) 519–526
11. Oh, J.Y., Stuerzlinger, W.: Moving objects with 2D input devices in CAD systems and Desktop Virtual Environments. In: Proceedings of graphics interface (GI), Canadian Human-Computer Communications Society (2005) 195–202
12. Masliah, M.R., Milgram, P.: Measuring the allocation of control in a 6 degree-of-freedom docking experiment. In: ACM Conference on Human Factors in Computing Systems (CHI), New York, NY, USA, ACM (2000) 25–32
13. Zhai, S.: User performance in relation to 3D input device design. *ACM Computer Graphics* 32(4) (1998) 50–54
14. Zhai, S., Buxton, W., Milgram, P.: The “Silk Cursor”: investigating transparency for 3D target acquisition. In: ACM Conference on Human Factors in Computing Systems (CHI), New York, NY, USA, ACM (1994) 459–464
15. Cacioppo, J., Tassinary, L., Berntson, G.: *Handbook of Psychophysiology*. Cambridge University Press (2007)
16. Picard, R., Vyzas, E., Healey, J.: Toward machine emotional intelligence: analysis of affective physiological state. *IEEE Trans. on Pattern Analysis and Machine Intelligence* (2001)
17. Shapiro, D., Jamner, L., Lane, J., Light, K., Myrtek, M., Sawada, Y., Steptoe, A.: Blood pressure publication guidelines. *Psychophysiology* (1996)