HAL
open science

# Towards Unification for Pointing Task Evaluation in 3D Desktop Virtual Environment

Mathieu Raynal, Emmanuel Dubois, Bénédicte Schmitt

## To cite this version:

Mathieu Raynal, Emmanuel Dubois, Bénédicte Schmitt. Towards Unification for Pointing Task Evaluation in 3D Desktop Virtual Environment. 1st International Conference on Human Factors in Computing and Informatics (SouthCHI 2013), University of Maribor; ACM Slovenia and IEEE Slovenia, Jul 2013, Maribor, Slovenia. pp.562-580, 10.1007/978-3-642-39062-3_35 . hal-01231767

**HAL Id: hal-01231767**

**https://hal.science/hal-01231767v1**

Submitted on 20 Nov 2015

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Open Archive TOULOUSE Archive Ouverte (OATAO)
OATAO is an open access repository that collects the work of Toulouse researchers and makes it freely available over the web where possible.

This is an author-deposited version published in : http://oatao.univ-toulouse.fr/
Eprints ID : 12580

The contribution was presented at SouthCHI 2013

Official URL: http://dx.doi.org/10.1007/978-3-642-39062-3_35

**To cite this version** : Raynal, Mathieu and Dubois, Emmanuel and Schmitt, Bénédicte *Towards Unification for Pointing Task Evaluation in 3D Desktop Virtual Environment.* (2013) In: 1st International Conference on Human Factors in Computing & Informatics (SouthCHI 2013), 1 July 2013 - 3 July 2013 (Maribor, Slovenia).

Any correspondance concerning this service should be sent to the repository administrator: staff-oatao@listes-diff.inp-toulouse.fr

# Towards Unification for Pointing Task Evaluation in 3D Desktop Virtual Environment

Mathieu Raynal, Emmanuel Dubois, and Bénédicte Schmitt

IRIT – ELIPSE, Université Paul Sabatier,
31062 Toulouse cedex 9, France
{firstname.name}@irit.fr

**Abstract.** New visualization systems for large and complex datasets are emerging and 3D Virtual Environments turn out to be a relevant solution. Interaction tasks in these 3D VE have been defined, especially to support evaluation of these applications. Nevertheless there is a lack of unified protocol to assess these elementary tasks in this context. Moreover it can be complex to determine the appropriate technique to perform these tasks as there is a lack of reference data. A standard is available for 2D pointing task, but there is no equivalence in 3D. In this paper, we propose an adaptation of this standard to a pointing task in a 3D VE. We detail our protocol and an instrumentation, which aims at assessing performance, comfort of techniques and satisfaction of users. We also present results of a user experimentation conducted according to this standard's adaptation.

**Keywords:** Pointing task, 3D virtual environment, 3D interaction, usability study.

## 1 Introduction

3D Virtual Environments (3D VE) are emerging in many different domains ranging from games to industrial applications (Geographic Information System - GIS, Computer-Aided Design - CAD, etc). A 3D VE can be defined as a space where one or more users interact in real time through physical means, and devices, with 3D digital data generated by a computer [1]. Recently new polyvalent technologies, like the Wiimote, successfully appeared in public spaces to offer advanced interaction techniques for using 3D [2]. They are initially intended to be used in game applications, and they should turn out to be satisfying and comfortable. In parallel, dedicated devices (e.g. Rockin'Mouse [3], Cubic Mouse [4]) have been proposed to offer advanced forms of interaction with 3D environments in more professional contexts. But they are uncommon and not yet part of the democratized and the well accepted set of interaction devices for professional activities; classical devices such as mouse or joystick remain the most common and used devices, even if they are not always well adapted to the tasks. Anyway to develop the use of advanced interaction techniques in

such professional contexts, these techniques should match several usability requirements: accuracy, execution speed and ease of learning. Moreover satisfaction and level of comfort should also be part of their acceptance due to their daily use. Therefore determining the device that can be efficient and well-accepted is a major need to contribute to the development of these advanced interaction techniques in such context, their evaluation and validation. It is also a complex need that requires appropriate means to measure and to compare techniques for 3D interactions.

So far, studies traditionally deal with one or several of the well established 3D tasks: selection, manipulation and navigation [5]. It results in complex user experiments, attempting to combine 6DOF input devices, 3D tasks, different interaction techniques, etc. And yet, focusing on a 3D pointing task only, which is a prevalent task, would simplify experiments' protocol. The pointing task consists in moving a cursor to a specific location by changing the coordinates of the cursor in translation (tx, ty, tz). A parallel can be drawn with the specification of an object position of the manipulation task and recommendations extracted from the evaluation of the pointing task could be transposed to this part of the manipulation task. Moreover a study of Masliah [6] revealed the interest to break the evaluation of the manipulation task into a separate evaluation of the translation and rotation tasks because users consider them as two different tasks. This reinforces the need to assess a low-level task interaction, such as the pointing task in 3D. However, even in such simplified contexts, facing the growing interest in advanced forms of interaction for 3D VE requires additional considerations: there is a need to support reproducible evaluations, in different application contexts, with a wide variety of advanced forms of interaction techniques; it is also required to support a composite approach, combining quantitative aspects, i.e. performance, and qualitative aspects, i.e. satisfaction.

To create a structured and replicable evaluation method for assessing pointing task in 3D Virtual Environments, we propose an adaptation of the ISO 9241-9 standard. Initially dedicated to pointing in 1D or 2D, this standard gives recommendations for assessing non keyboard technique and combines quantitative and usability measures. Usability here refers essentially to effort and comfort. Our adaptation aims at extending its applicability to a 3D environment and to additional consideration related to the performance and the user satisfaction. After reviewing existing works for evaluating interaction techniques in 3D environments, we describe and justify the required adaptations in the ISO standard. We finally present a software tool used to conduct a user's evaluation of a 3D pointing task and report on the evaluation conducted on the basis of the adaptation and its associated tool to compare and assess two different interaction techniques for pointing in 3D.

## 2 Evaluation of Input Devices in 3D Environments

Two approaches presented in the literature to assess input devices in 3D environments can be opposed: structured approaches for pointing task evaluation and ad-hoc approaches for 6-DOF interaction techniques evaluation.

## 2.1 Pointing Task Evaluation Approaches

**Fitts' Law and its Extensions.** Fitts' Law is originally designed to predict the time (MT) required to move to a target in 1D where a and b are two constants, determined empirically for a device and an interaction context. W is the width of the target and D the distance to the target. The log term is the Index of Difficulty (ID) and is measured in bits.

$$MT = a + b \times ID \text{ with } ID = \log_2(D/W + 1) \tag{1}$$

Recently, some models have been proposed to extend Fitts' Law. MacKenzie and Buxton [7] proposed different approaches to compute the index of difficulty in 2D. But Zhai and Accot demonstrated some limitations to those models, especially when target height and target width are not equal, and proposed in reaction a model for bivariate pointing.

Ware and Lowther [8] proposed an extension of the Fitts' Law for 3D. Targets are now defined in 3D with a width, height and depth and the computation of the indexes of difficulty relies on the smallest of these three dimensions. However this extension of the model has the same issues as the reference model. Furthermore according to our knowledge, no experience with explicit 3D targets has already been performed to fully validate this model. Grossman and Balakrishnan [9] adaptation consists in studying physical movements of users in a XZ plane for a volumetric display. They validated their model for a volumetric display but not in a 3D Virtual Environment. Moreover, movements were constrained on one plane (X,Z) with (Y=0), so movements were not evaluated on all axes. So far these alternate models have not overcome the classical Fitts' Law to study and predict pointing performance in a 3D VE. Hence the classical model is used regardless of the dimensions.

**Throughput.** The throughput (TP) has been introduced by the ISO 9241-9 standard to describe and compare the performance of at least two pointing devices in a given context. It is measured in bits per second and the formula to compute the TP is:

$$TP = \frac{ID_e}{MT} \text{ with } ID_e = \log_2(\frac{D_e}{W_e} + 1) \tag{2}$$

This is an overall performance measure, based on the Fitts' model, and is independent of the speed-accuracy tradeoff contrary to the Index of Performance. MT is the average movement time performed by users over a block of trials and IDe is the effective Index of Difficulty. These terms represent how a user effectively performed rather than what the user can do in theory. Indeed De is the effective travelled distance between two targets covered by users during the evaluation and We is the effective width corresponding to the distribution of users' selections in or around the target during the evaluation. But no throughput measures are currently available for 3D interaction techniques in 3D VE.

**ISO 9241-9 Standard.** The ISO 9241-9 standard [10] gives requirements about design and evaluation of non keyboard devices such as mouse, joystick, etc. The standard provides recommendations and a set of variables for different kind of tests depending on the task to evaluate: for example tapping and dragging test respectively corresponds to a pointing and a drag-and-drop task. The use of this standard aims at simultaneously measuring the throughput of such devices, to compare the performance of devices, and leading a qualitative study based on a comfort assessment and an effort scale. This standard has been used in recent studies to assess the performance of diverse techniques, e.g. [11]. However these recommendations concern 2-DOF interaction techniques used to perform 1D or 2D tasks. The standard does not give recommendations to assess interaction techniques with 3-DOFs or more and does not address 3D environment. Nevertheless some attempts have been proposed to extend the ISO standard to a 3D environment.

**Proposition of Adaptation of the ISO 9241-9 Standard to a 3D Environment.** Recent studies of Teather et al. [12,13] deal with pointing task evaluation in a 3D environment based on the ISO standard. In the first study [12], targets are spheres disposed on a 2D circle on a vertical plane. In the second study, the representation used is a 3D scene [13] in which targets are circles placed on cylinders. In both cases, due to the perspective rendering mode, targets representation appears on a plane, either horizontal or vertical: selecting a target thus results of the combination of a 2D pointing on the plane with the use of a ray-casting technique to reach the appropriate depth. In these settings interactions with the device are thus confined to the plane (X,Y): the depth is not processed with one specific DOF but is a software computation. Cursor movements are therefore not considered in all directions. Moreover these studies take place in immersive environments and not desktop environments.

Another recent study [14] evaluates 3D object manipulation using virtual hand techniques. They based their evaluation on the ISO standard and adapted the formula to compute a 3D projection, onto the task axis, of the user's selection, i.e. of the position of the cursor when the user validated the target selection. Nevertheless their evaluation took place in a Tangible Augmented Reality Environment and they did not propose a digital 3D VE to interact with. This work illustrates the interest to base an evaluation on the ISO standard but their results can not be compared with evaluations conducted in 3D Virtual Environment like Teather's works.

Beyond these propositions, we are not aware of any other study based on the ISO standard performed in a 3D VE and involving interaction techniques with 3DOF or more. Therefore results provided by structured approaches do not yet constitute reusable references. More concrete results have been measured through empirical studies. The following section summarizes ad hoc evaluations lead in 3D Virtual Environments.

## 2.2 Evaluation of 3D Interaction Techniques in 3D Environments

Among existing studies in the domain of 3D interaction techniques evaluation, we here focus on three representative examples: (1) the Airmouse technique [15], (2) a study lead by Berard [16], (3) a study lead by Zhai [17].

These studies differ in terms of task and settings. In terms of settings, the studies (1) and (2) manipulate only 3 axes for translations while the study (3) uses all 6 axes for translations and rotations. In terms of task, study (1) assesses a pointing task: moving a little sphere in a spherical area; study (2) evaluates a placement task: moving a cube as close as possible to a target represented by a cube with a smaller size; study (3) evaluates a docking task: manipulating a tetrahedron in another tetrahedron with the same size. This task consists in putting the object exactly in the same position and orientation than the target. But for each task, users aim at moving a cursor or an object as quickly as possible to a specific location.

Even if all these experiments focus on quantitative results, in particular performance time, each one considers and relies on different variables. Studies (2) and (3) define an error rate. Studies (1) and (2) use the notion of index of difficulties, as defined in Fitts' law to modify task conditions, but only study (1) matches indexes of difficulties and times to perform the task. However traditional parameters a and b of a Fitts' law analysis are not studied. In addition, as opposed to study (2), studies (1) and (3) include qualitative measurements by assessing the fatigue; however the evaluation of the fatigue is not similar among the two studies.

These studies show a lack of formalization for 3D interaction techniques evaluation and highlight limits that are commonly observed in similar users' studies: DOF considered, task, measurements and method. Furthermore the analyzed measures are not identical and there is no overall measure to easily compare these interaction techniques.

To summarize, the first set of approaches gives recommendations to assess the pointing task in a structured way, while the second set of approaches provides concrete results and measures, which are complex to compare. Our goal is thus to take advantage of the structured approach to guide and to standardize the design of such experiments. We also take advantage of the variables and measures promoted in the ad-hoc approaches and denoting relevant aspects of the users' experiment in 3D VE. Our work is based on an extension of the ISO standard.

## 3 Unification for the Pointing Task in a 3D Desktop VE

Among the type of tests covered by the standard, we focus on the multi-directional tapping test as we intend to evaluate the pointing and selecting tasks. The movement is multidirectional as we consider pointing movements in any direction (x, y and z). The adaptation we propose is explicitly targeted at pointing tasks in 3D context. A first introduction to this adaption has been briefly presented in [18]. In this section, we more clearly refer to the sections of the existing standard for which modifications are proposed. We therefore detail 5 main dimensions addressed by the initial standard, three of them related to the performance (task, targets, and variables) and two related to qualitative aspects (representation and finally, comfort and satisfaction). For each one, we present and justify the adaptation introduced to fit with a 3D pointing context.

### 3.1 Adaptation from a Performance Point of View

Three major aspects of the standard structure the quantitative assessment: task, targets and variables.

**Task.** In the ISO standard, the recommended task consists in selecting a fixed target as quickly and accurately as possible (ISO 9241-9 - Annex B.B.3). Selection comprises the pointing of a target by moving a cursor from a starting target to a destination target with the pointing technique. The pointing task ends in the validation of the selection when confirming the selection by pressing a button or key. If the cursor is inside the destination target when the user confirms the selection, the pointing task is successful. If the cursor is outside the target, the validation raises an error. In every case, pressing the button or key is the end of the current task and the beginning of the next pointing task of the protocol. The task difficulty varies according to different IDs as defined in Fitts' Law: targets size and distance between targets are modified over the experiment.

In our adaptation, the task instruction, the existence of a source and destination target and the need for a confirmation has been kept. The adaptation first concerns the validation. Indeed, the 3D cursor does not need to be fully inside the target to consider the validation as a successful selection: when validating, detecting a collision between the target and the cursor is sufficient. In the absence of collision, the validation raises an error.

The ISO standard also describes the use of the task axis (line between the start target and the end target) implied by user's movements during the task and used to measure the throughput: this has been kept in our adaptation.

**Targets.** Although the form of targets is not explicitly constrained by the standard, circle targets are the most used in evaluations of the literature, because a circle has one width regardless of the user's movement. As a result, our adaptation to 3D recommends that targets become spheres to keep one measure for the width regardless of the 3D movement: the size of the target will thus be given by the diameter of the spheres.

Regarding the arrangement and position of targets, the ISO standard recommends to equally space the targets around the circumference of a support circle (ISO 9241-9 - B.B.6.2.2). In addition, the place of the target on the circle must be defined so that each target has a symmetric with respect to the center of the support circle. This ensures that the distance between two targets is always the same, and equals to the support circle diameter (ISO 9241-9 - B.B.6.2.2).

In our adaptation, the arrangement of the targets must be revisited to fit with 3D: the support circle is thus replaced by a support sphere. Targets are still equally spaced on the circumference of a support sphere. But their position is defined so that each target has a symmetric with respect to the center of the support sphere. And the distance between two targets is still equal to the diameter of the support sphere and constant over the task.

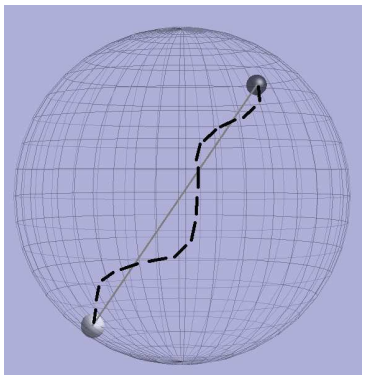
**Variables.** The ISO standard recommends different independent and dependent variables (ISO 9241-9 - B.B.5). Independent variables inherent to the experiment are input technique, targets size, targets distance and the resulting IDs. Dependent variables mentioned in the ISO standard include the throughput and the mean movement time.

In our adaptation, these variables have been kept. However the calculation of IDs is modified, because the cursor is volumetric. Its computation, as illustrated in the equation below, thus includes the size of the cursor:

$$ID = \log_2\left({}^{D}/_{W_{target} + W_{cursor}} + 1\right) \tag{3}$$

The calculation of the effective IDs (IDe) is also modified and relies on the coordinates of the point effectively reached by the center of the 3D cursor when the user selects the target. The effective width is then computed by solving a 3D parametric equation to project the effective users' selection point onto the "task axis". The effective distance corresponds to the 3D Euclidean distance computed between the starting users' selection point and the destination users' selection point of the "task axis". The effective width and distance, used in the computation of the TP, are computed as an average over all repetitions for a given condition.

Our adaptation of the standard also recommends considering two additional dependent variables: the error rate and the inefficiency ratio. Very frequently used as a dependent variable in ad-hoc approaches, we recommend considering the error rate because it complements the TP measure. Furthermore the 3D context induces more movements' variability during the task than 2D contexts; a deeper investigation of the users' performance is thus required to qualify this quantitative measure. Masliah [19] suggests the inefficiency ratio, defined by Zhai and Milgram [20], to support the analysis of the performance quality: this ratio compares the user's travelled distance to the optimal distance (cf. Figure 1). This ratio indicates the coordination of users during the task: the closer the ratio is to 0, the better is the coordination as 0 indicates that users navigated the optimal distance. The inefficiency ratio therefore complements the error rate, one aspect of users' performance. We thus recommend considering this variable in the set of dependent variables.



$$Inefficiency\ ratio = {}^{UD - OD}/_{OD}$$

Dashes trajectory : User's travelled distance (UD) = 309px
Straight trajectory: optimal distance (OD) = 150px
IR = (309 – 150) / 150 = 1.06

**Fig. 1.** Illustration of the coordination for one task

### 3.2 Adaptation from a Qualitative Point of View

Two elements structure the qualitative assessment: representation with especially 3D indices and, comfort and satisfaction.

**Representation and 3D Indices.** The ISO standard does not recommend any specific elements of representation. Despite that, the pointer is commonly represented by a cross. In our adaptation, the volumetric representation of the cursor we recommend is a red sphere. In addition its recommended width is about 70% of the littlest target width (10px in our case) to remain visible regardless of its position in the 3D environment. The shape of the cursor is defined in order to ensure that the collision point is captured at a same distance from the cursor center to the target center whatever the 3D movement, which is not possible with a square.

Our adaptation to 3D VE is also facing the traditional problem of depth perception in 3D. Due to the desktop environment and the lack of effective hardware and software solutions, users may not be able to correctly sense the depth of a digital object, thus making the modeling of the 3D pointing relatively difficult. For these reasons, 3D perception requires to define some additional and specific aspects to the representation. Some 3D visual indices are defined to overcome difficulties of depth perception: a color graduation combined with a transparency [21] of the target spheres. This transparency reduces the occlusion with other spheres and allows users to see the cursor behind spheres. A relative size of spheres is due to the perspective of the 3D scene, so spheres placed behind seem to be smaller. Furthermore, the support sphere is represented by a transparent wireframe. These indices have been explored and defined during an iterative design made with the participation of 3D experts. Although shadows are commonly used in 3D scenes, our adaptation does not encourage the use of shadows to avoid overloading the environment. Indeed Hubona et al. [14] showed that user performance for a docking task decrease when increasing shadows number and scene complexity. These recommendations thus constitute a scene of reference for evaluating interaction techniques for pointing tasks in 3D; using this common scene of reference will avoid inserting biases from the depth perception that could jeopardize evaluation results of interaction in 3D spaces.

**Comfort and Satisfaction.** The ISO standard recommends two complementary rating scales to assess the comfort of interaction techniques: a comfort questionnaire to measure the comfort and fatigue of techniques and an effort scale to measure the effort perceived by users during the task (ISO 9241-9 - Annex C).

In our adaptation, the two rating scales are kept and we reinforce the usability analysis of the comfort questionnaire by adding the SUS questionnaire that covers user's satisfaction and learning. Indeed the comfort questionnaire does not cover aspects of usability, such as user' satisfaction, learning and users' preference. SUS includes 10 items and computes an overall usability score for each technique, useful to easily compare different techniques assessed in a same context. We further enrich the analysis of the satisfaction aspect by inserting the use of 6 SUMI items focusing on the frustration and the stimulation for example. Finally we complete this part of this

qualitative aspect by asking users their 3 preferred points and their worst points about each 3D input technique.

In addition to these main adaptations of the ISO standard, we developed an interactive tool to support the use of the resulting adaptation when performing an evaluation. Its main features are presented in the next section.

### 3.3 Instrumentation

The interactive tool aims at supporting the deployment of the protocol to evaluate 3D input techniques and at collecting logs. The interactive tool is composed of a 3D environment (cf. Figure 2) and a configuration panel. The 3D graphics software is written in C++ using OSG toolkit. The configuration panel is written in C#, and to propose an evolutive software tool, all software elements are linked through a software bus called Ivy. The pointing task is composed of a number of targets that the experimenter can manage with the configuration panel: targets can be placed at specific longitude and latitude. A block is composed of a series of pointing tasks, which have the same Fitts'ID. The experimenter can configure the targets' size (width) and the support's size (distance) to adjust the ID of the block. Running the task on one block is a three-step process: 1) the pointing task begins when the participant selects the starting sphere, colored in blue, without making an error. 2) It then highlights the symmetric destination sphere, colored in yellow. 3) The task ends when the participant selects the destination sphere, even if there is an error (i.e. no collision with the cursor). The target sphere becomes orange when the cursor and the target are in contact.

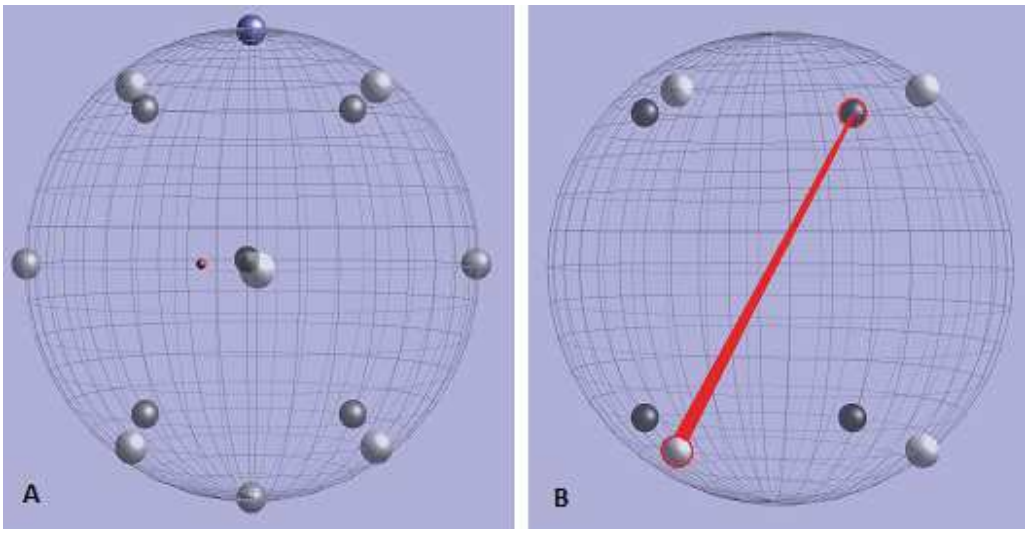



**Fig. 2.** Illustration of the 3D scene (A), of the depth perception indices and of a task axis (B)

During the task, the software tool records temporal data, such as beginning and end of each pointing task, validation of the selection and execution time. It logs the position of the 3D cursor and errors when the validation was done without collision between the cursor and the target. All experiment events are dated and recorded in a XML format.

In order to concretely show the use of the recommendations of our adaptation, we present a case study, the protocol for the evaluation of two interaction techniques. We then analyze and discuss results of this evaluation.

# 4 Concrete Use of the Adaptation of the Standard and Evaluation Results

The case study we consider in this section is taken from an industrial context. A partial set of results regarding has already been mentioned in [22, 23].

Nevertheless, instead of just using an ad hoc evaluation protocol as in these previous publications, we here tightly follow the adaptation of the standard presented in this paper to measure and analyse a well identified and justified set of variables and protocol steps.. We first detail the experimental protocol and report on the quantitative and qualitative results of this evaluation.

## 4.1 Case Study: Aircraft Static Test Design

Our case study is based on a real application proposed to design and manage campaign of static test for aeronautics (cf. Fig. 3). Designers have to manipulate and position sensors on a 3D digital mock-up, which represents a real aircraft, including dimensions and volume. Interactions with these models enable users to perform tasks in 3D as if they were manipulating the real aircraft. Users can apply transparency on parts of the digital mock-up to easily interact with otherwise hidden elements of the 3D structure: techniques like ray-casting are not usable given the digital mock-up complexity, and its multi-layered structure.

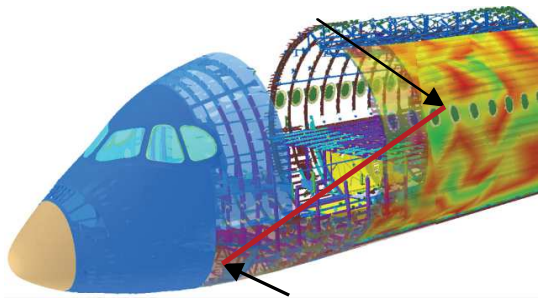


**Fig. 3.** : View of the 3D model in EasySensorManager application

**Task.** 3D pointing task is regularly performed to place or manipulate sensors in the 3D aircraft. Designers frequently have to successively select two elements of the digital mock-up to measure a distance for example: this corresponds to the selection task described in the standard between two points of the aircraft placed at different depth in the mockup (cf. Fig. 3 - points A and B).

**Targets.** Targets are typically elements of the plane structure (e.g. extremities of wings) or sensors. Designers can select sensors placed on the inside or outside surface

of the nose plane, the left or right wing and the tail plane. The volume of these elements and the distance between them are variable, and these variations match different task difficulties (ID) of the standard.

**Variables.** Designers have to accurately position a large number of sensors for one test, so small movement times and error rates are important to reduce the time of the task.

**Representation and 3D Indices.** To help perceive the 3D in the complex mock-up and to avoid jeopardizing its understandability, no "decorating" elements, such as shadows, are added to the digital mock-up. Conversely objects are represented in different colors and with different transparencies to highlight different depths present in the digital mock-up. This is also in line with the recommendation expressed by the standard.

**Comfort and Satisfaction.** The sensors positioning is a daily activity for designers and they can spend many hours to realize this activity; it is therefore important that they feel comfortable and satisfied by the techniques they use. SUS score and other fatigue scale will be helpful in this use case.

Two techniques have been designed and implemented to support this activity. Concretely, the first technique is an advanced interaction technique: a physical object is manipulated in a 3D physical space in order to move the digital pointer in the 3D mock-up. The position of the object in the physical space matches with the position of the pointer in the 3D environment. This technique is based on a position-control mode and constitutes an isotonic controller. The second technique is based on the users force applied to an object similar to a joystick in order to move the digital pointer. The input force is transformed into speed to change the position of the pointer in the 3D environment. This technique is based on a rate-control mode and corresponds to an isometric controller. The evaluation of the pointing task with these two interaction techniques will illustrate the use of the standard adaptation.

### 4.2 Participants

Fourteen non-paid users (3 female and 11 male) were recruited to participate in this study and to test the two techniques. But usable results were only available for 13 users, because logs for one user were corrupted due to an issue with data recording. Participant ages ranged from 22 to 39 (mean: 27.5, SD: 5.26). They were regular computer users but they were not familiar with our 3D interaction techniques. They used their dominant hand (12 right-handed and 1 left-handed participants).

### 4.3 Hypothesis of the Experimentation

A goal of the evaluation is to assess the speed and the accuracy of techniques, our second hypothesis (H1) is that the isotonic technique is more efficient than the

isometric technique and has a better throughput measure. We also hypothesize that users prefer interacting in a 3D environment with the isotonic technique (H2) as the interaction seems more usual than for the isometric technique. The assessment of these hypotheses will rely on the results provided by our adaptation of the standard.

### 4.4 Apparatus

The experiment was conducted using a HP EliteBook 6930p laptop PC with Windows 7 and a LCD screen with a resolution of 1920*1080. The computer's processor was a 2.4GHz Intel Core Duo CPU with 4GB of RAM and an ATI Mobility Radeon HD 3450 graphics controller of 256Mo. The goal was to compare two interaction techniques presented above: the technique based on a position-control mode and the one based on a rate-control mode were respectively implemented with a Polhemus Patriot, an isotonic tracking system, and the SpaceNavigator, an isometric joystick. In order to implement these interaction techniques, a pre-experiment has been conducted to determine optimal parameters for position-control and rate-control modes.

### 4.5 Procedure

At the beginning, each participant completed a consent form and a demographic questionnaire. Then participants sat in front of a desk where the LCD screen and both devices were placed. Instructions, in accordance with those specified in the adaptation of the standard, were given to participants and they performed a 15 minutes-training before the experiment. As mentioned in the adaptation, the pointing task began when participants selected the starting sphere. The task ended when participants selected the destination sphere. Participants were allowed to rest whenever they wanted between two pointing tasks. When all pointing tasks were performed with the first technique, they filled out a scale to assess technique comfort and a satisfaction questionnaire to assess technique usability. Participants followed the same procedure for the second technique. At the end of the experiment, participants filled out a questionnaire in order to compare both techniques. Complementary interviews, as recommended in our adaptation, were used to gather users' preferences and to identify the three best and the three worst points of each technique. In this case study, the effort scale was not used as it is generally used to assess large muscle groups (arm, shoulder and neck) but may not be appropriate for smaller muscles, such as those involved in manipulating our two techniques. Assessment of efforts was thus limited to the use of the 5 indices of the comfort scale.

### 4.6 Design

As recommended in our adaptation, different IDs need to be specified. In our experiment, the IDs were defined as a combination of width and distance. Unlike [23], we chose to present, in this article, indices of difficulty according to the calculation

presented in (3). Cursor size is 10 px so target width + cursor width=25px, 40px, 55px and distance=150px, 450px, 750px. IDs were ranged from 1.898 to 4.954. In this instance of the adapted standard, the number of targets to point has been fixed to 14 targets: 6 targets were placed on the extremities of the axes and 8 at longitude and latitude Π/4. Participants were divided into 2 groups and each group tested techniques in a counterbalanced order.

Finally, we collected 3276 pointing times as we had 13 participants * 2 devices * 9 blocks (counterbalanced by a Latin square) * 14 target selections.

### 4.7 Quantitative Measurements

Based on the logged data, we analyze the quantitative results by use of descriptive statistics and inferential statistics (i.e. One-way RM ANOVA).

**Movement Time.** The ISO standard includes the mean movement time as a dependent variable based on the Fitts' law. So we analyzed the mean time required to select a target: the mean time measure for the isometric technique (hereafter referred to as the SpaceNavigator) is 10681ms and for the isotonic technique (hereafter referred to as the Patriot) is 6265ms. The analysis shows a significant difference between these techniques ($F(1,12) = 179.05$, $p < 0.0001$). So the Patriot enables users to perform the pointing task significantly faster than the SpaceNavigator. Based on observation and answers during the semi-guided interviews, we believe that this higher efficiency of the Patriot is due to the use of more usual and less constrained arm motions.

**Throughput.** Among the dependant variables identified in the standard, the throughput is a performance measure to easily compare two techniques. Computing the throughput for each input technique relies on effective widths and effective distances. This computation uses a per-participant and a per-block basis to produce the "grand throughput". Throughputs for the isotonic technique and the isometric technique respectively are 2.1bps and 1.69bps, that is to say the rate of information transfer during the task. So we infer that the Patriot is more efficient than the SpaceNavigator. This confirms our hypothesis (H1) about the higher efficiency of the isotonic technique.

**Inefficiency Ratio and Movements Analysis.** As a complement to the TP, the measure of the inefficiency ratio is recommended in our adaptation of the variables of the standard. The computation of the inefficiency ratio for both techniques reveals that the translation inefficiency of the isotonic technique (1.06) is lower than the translation inefficiency of the isometric technique (1.56). This means that translation trajectories with the isotonic technique are 106% longer than the optimal distance and translation trajectories with the isometric technique are 156% longer than the optimal distance. The analysis shows a significant difference between these techniques ($F_{(1,12)} = 218.75$, $p<.0001$). We infer that the isotonic technique supports a better coordination than the isometric technique.

In addition to the quantitative measures recommended in the adaptation of the standard, we analyzed users' movements to better understand the impact of the 3D on results of this experiment. Movement required to achieve every pointing task is performed in 3D. We analyzed the difference between movements involving translations realized along one axis (simple movements) and movements involving a combination of translations along the 3 axes (complex movements). A first analysis shows a significant difference between movements in terms of time ($F_{(1,12)}$ = 13.67, p<0.001) and distance ($F_{(1,12)}$ = 16.91, p<0.0001) regardless of the technique (see Fig. 4 and Fig. 5). Users perform the task quicker and with a shorter travelled distance when simple movements are involved than when complex movements are proposed. This means that users have less difficulty to manipulate the pointer during simple movements than during complex movements.

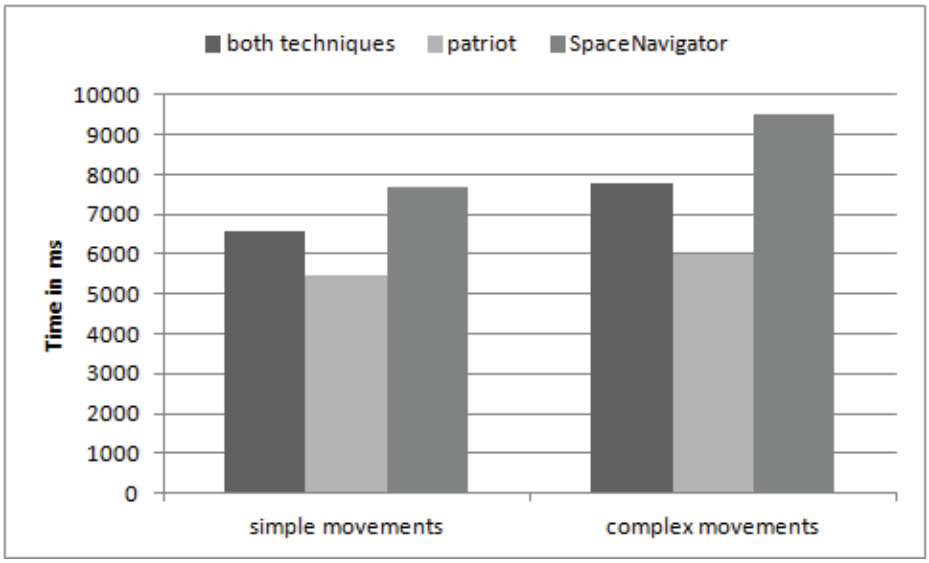



**Fig. 4.** Time to complete the task according to the difficulty of the movement and to the device

A second analysis shows a significant difference between the two techniques in terms of time ($F_{(1,12)}$=56.74, p<0.0001) and of distance ($F_{(1,12)}$=4.81, p<0.05) with simple movements (see Fig. 4 and Fig. 5). A significant difference also exists between the two techniques for the time ($F_{(1,12)}$=123.86, p<0.0001) and for the distance ($F_{(1,12)}$=47.68, p<0.0001) during complex movements. The *Patriot* enables users to perform the task quicker and with a shorter travelled distance for both kinds of movements. These results confirm the inefficiency ratio and the best performances of the isotonic technique.

**Error Rate.** Our adaptation of the standard establishes the measure of the error rate as a dependent variable. So we analyzed error rate for both techniques to know how successful the pointing task is and to determine the effectiveness of these techniques. The measured error rate of the isometric technique is lower than the one of the

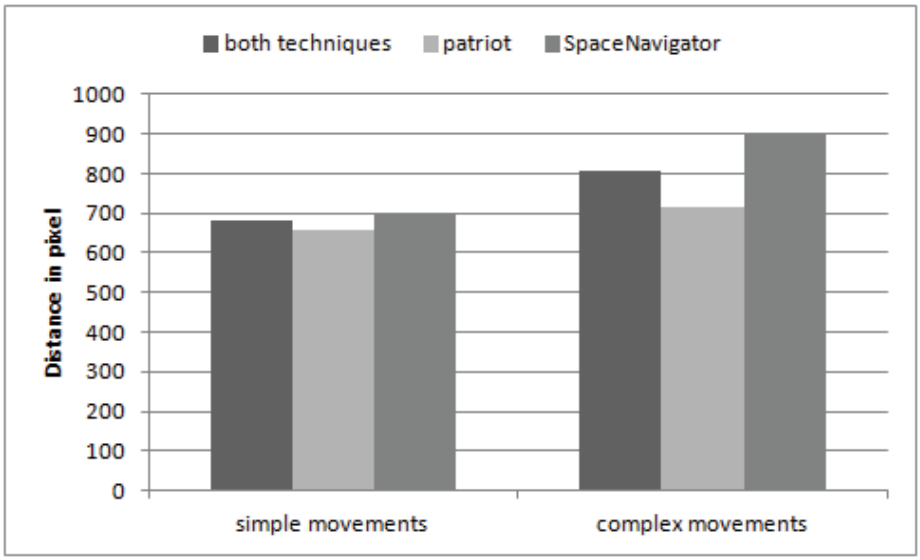



**Fig. 5.** Distance traveled by the cursor to complete the task according to the difficulty of the movement and to the device

isotonic technique (4.4% vs 7.6%), which can explain the difference of mean movement times. The analysis shows a significant difference between these techniques ($F(1,12)=20.77$, $p<0.005$). In-study observations highlight that the SpaceNavigator enables users to realize small motions in order to precisely select targets and to maintain a neutral position as opposed to the Patriot, for which users have difficulties to keep the pointer in contact with targets.

### 4.8 Qualitative Measurements

Our composite approach also includes qualitative aspects that we analyzed by use of non-parametric statistics.

**Comfort Assessment.** Participants completed the comfort questionnaire, which includes two indices, general indices and fatigue indices, and based on a five-level Likert scale. As described in the procedure, in a first phase, participants completed a questionnaire after testing the first technique in order to assess the comfort of this first technique. Results of the first questionnaire do not show any major significant results for all indices according to the Mann and Whitney test, except for the arm fatigue. The mean score over all scales of the first phase is 3.41 (SD=0.45) for the isometric technique and 3.54 (SD=0.35) for the isotonic technique: this result shows that users tend to have a better feeling about the comfort of the isotonic technique. After testing the second technique in a second phase, participants filled out a scale to compare both techniques according to the same indices. A sign test shows no significant difference ($|Z|=-0.28$, $p>0.5$) between both techniques. But results of the two phases ($p=0.091$ for the first phase) reveal that the isotonic technique tends to cause more fatigue than the isometric technique for arm or shoulder, so it would be relevant to envision another design for this point of focalization.

**Users' Satisfaction.** The adaptation of the standard recommends analyzing usabilit of techniques with a SUS questionnaire in addition to the comfort questionnaire. W computed a SUS score to measure the usability of each technique. The score of th isometric technique (52.5) is marginal at a low level whereas the score of the isotoni technique (76.3) is acceptable, according to [24]. The techniques' usability can b improved, even if the usability of the isotonic technique is good. Both SUS score establish that the current usability of the techniques does not jeopardize their use fc pointing tasks. Besides, answers to the SUMI items questionnaire highlight the isotor ic technique is perceived as more stimulating and less frustrated than the isometri technique.

**Users' Preferences.** As proposed by the adaptation of the standard, we used inte views since users' answers are usually a fair indication of improvements, indicatin what could be developed to reach a better usability level. Users mentioned the mos frequently a good aestheticism of the device, as the best point of the isometric tech nique (based on the *SpaceNavigator*) and an inappropriate association between inpu and output interactions, especially depth interactions, as the worst point of this tech nique. They mentioned the most frequently the ease of use, as the best point of th isotonic technique (based on the *Patriot*), and the fatigue lead by the technique durin the task, as the worst point. Results of the interviews highlight a main differenc about these techniques: users did not well perceived the depth by manipulating th isometric technique whereas they could find bearings in space with the isotonic tech nique to move easily in 3D and it seems to be more "intuitive". Besides, users mad no negative comments about the representation and the 3D indices suggested in ou standard's adaptation.

With regards to users' preferences, 85% of users like to interact with the isotoni technique and 77% of them estimate this technique is more usable than the isometri technique. This confirms our second hypothesis (H2). Indeed they are 69% to esti mate that the isometric technique is more constraining in terms of freedom c movement.

## 5 Discussion

This paper progressively introduces and justifies the modifications required to adjust standard for evaluating 2D pointing to the specific context of 3D VE. The adaptatio of the standard includes 3 main aspects of enrichment. The first aspect is the additio of the SUS questionnaire to the comfort questionnaire to enable an easy compariso between techniques in terms of usability; it also provides an overall measure, simila to the throughput but related to satisfaction and not to performance. The second aspec is the set of variables of the experiment, which is enriched by a measure of coordinatio to better understand users' movements in 3D. The third aspect is the adjustments c the task and targets to fit with 3D pointing contexts: forms, arrangements, representa tions and 3D indices are proposed in order to help users better perceive the depth

We also propose a software tool to support the execution of experiments in a 3D VE in accordance to the adaptation of the standard.

As a result, our adaptation of the standard presents a generic evaluation method and proposes a structured protocol to lead the evaluation of interaction techniques for pointing task in 3D VE. According to Zhai [25], the usability of 3D input techniques can be defined by 6 criteria: speed, accuracy, fatigue, device persistence and acquisition, coordination and at last, ease of learning. This adaptation offers a composite approach and a set of recommendations, which address all the criteria defined by Zhai. Indeed, in our adaptation, speed and accuracy are measured by the throughput. The fatigue and, persistence and acquisition are evaluated with the comfort questionnaire from the standard and additional semi-directive interviews. The coordination is computed with the inefficiency ratio that we recommend in our adaptation. To measure the ease of learning, we recommend adding an overall score of usability and users' satisfaction measured thanks to the SUS questionnaire.

Furthermore, we show in this paper the use of our adaptation on a concrete case study. It reveals the advantage of having a generic and composite approach. Overall scores of throughput and SUS support a quick and overall comparison of 2 techniques: in our case study, it revealed that the isotonic technique is more efficient and usable than the isometric technique. Moreover, interviews show that users preferred using the isotonic technique to move the pointer as the manipulation seemed more natural, although it may infer more fatigue. This experiment also highlights a tradeoff between the mean movement time and the error rate: the confidence measure of the SUS questionnaire shows that users were more confident using the isotonic technique than the isometric technique and may explain that they paid less attention to the task accuracy. These results enable a complementary analysis to better understand the users' performance and satisfaction for a pointing task in 3D VE.

Using the adaptation of the protocol lead to some already known results (tradeoff between error and speed, preference towards isotonic device, etc), but also provided a framework to explore additional considerations and refine the initial outcomes (inefficiency ratio, impact of ID, complexity of movements, SUS question).

Being an extension of an existing standard, it contributes to the standardization of pointing evaluation in 3D VE. Future uses of the standard by designers will therefore contribute to classify the results according to considerations that are specific to 3D pointing and highlighted in this adapted protocol. Actually this protocol is a guide to set up an experiment and it will contribute to empirically generate recommendations based on these results.

## 6 Conclusion

In this paper we presented an adaptation of the existing ISO 9241-9 standard to a pointing task in 3D VE. We based our work on an existing standard addressing 2D situations only, because the literature in 3D mentions elements for 3D evaluation that are not present in the ISO 9241-9 standard and does not refer to other standards.

The adaptation of such a standard presents the advantage to propose a structured approach to evaluate a 3D low-level interaction task combining quantitative and qualitative aspects. It offers a unified method to perform pointing task evaluation in 3D VE and constitutes a first step towards 1) a standardization of such procedures and 2) the starting point for the creation of a coherent usability knowledge for pointing in 3D VE.

## References


1. Bach, C., Scapin, D.: Ergonomic criteria adapted to human virtual environment interaction. In: 15th French-Speaking Conference on HCI, pp. 24–31. ACM, New York (2003)
2. Gallo, L., de Pietro, G., Marra, I.: 3D interaction with volumetric medical data: experiencing the Wiimote. In: 1st International Conference on Ambient Media and Systems (AmbiSys 2008) (2008)
3. Balakrishnan, R., Baudel, T., Kurtenbach, G., Fitzmaurice, G.: The Rockin'Mouse: integral 3D manipulation on a plane. In: SIGCHI Conference on Human factors in computing systems (CHI 1997), pp. 311–318. ACM, New York (1997)
4. Froehlich, B., Plate, J.: The cubic mouse: a new device for three-dimensional input. In: Conference on Human Factors in Computing Systems (CHI 2000), pp. 526–531. ACM, New York (2000)
5. Bowman, D.A., Kruijff, E., LaViola, J.J., Poupyrev, I.: 3D User Interfaces: Theory and Practice. Addison Wesley Longman Publishing Co., Inc.
6. Masliah, M.: Measuring the allocation of control in 6 degree of freedom human-computer interaction tasks, PhD thesis, University of Toronto (2001)
7. MacKenzie, I.S., Buxton, W.: Extending Fitts' law to two-dimensional tasks. In: Conference on Human Factors in Computing Systems (CHI 1992), pp. 219–226. ACM, New York (1992)
8. Ware, C., Lowther, K.: Selection using a one-eyed cursor in a fish tank VR environment. ACM Trans. Comput.-Hum. Interact. 4(4), 309–322 (1997)
9. Grossman, T., Balakrishnan, R.: Pointing at trivariate targets in 3D environments. In: Conference on Human Factors in Computing Systems (CHI 2004), pp. 447–454. ACM, NY (2004)
10. ISO, 9421-9 Ergonomic requirements for office work with visual display terminals (VDTs)-Part 9: Requirements for non-keyboard input devices. ISO (2000)
11. Natapov, D., Castellucci, S.J., MacKenzie, I.S.: ISO 9241-9 evaluation of video game controllers. In: Graphics Interface 2009 (GI 2009), pp. 223–230 (2009)
12. Teather, R.J., Stuerzlinger, W.: Target Pointing in 3D User Interfaces. In: Poster at Graphics Interface (GI 2010) (2010)
13. Teather, R.J., Stuerzlinger, W.: Pointing at 3D targets in a stereo head-tracked virtual environment. In: IEEE Symposium on 3D User Interfaces (3DUI 2011), pp. 87–94 (2011)
14. Hubona, G.S., Wheeler, P.N., Shirah, G.W., Brandt, M.: The relative contributions of stereo, lighting, and background scenes in promoting 3D depth visualization. ACM Trans. Comput.-Hum. Interact. 6(3), 214–242 (1999)
15. Ortega, M., Nigay, L.: AirMouse: Finger Gesture for 2D and 3D Interaction. In: Gross, T., Gulliksen, J., Kotzé, P., Oestreicher, L., Palanque, P., Prates, R.O., Winckler, M. (eds.) INTERACT 2009. LNCS, vol. 5727, pp. 214–227. Springer, Heidelberg (2009)

16. Bérard, F., Ip, J., Benovoy, M., El-Shimy, D., Blum, J.R., Cooperstock, J.R.: Did "Minority Report" Get it Wrong? Superiority of the Mouse over 3D Input in a 3D Placement Task. In: Gross, T., Gulliksen, J., Kotzé, P., Oestreicher, L., Palanque, P., Prates, R.O., Winckler, M. (eds.) INTERACT 2009. LNCS, vol. 5727, pp. 400–414. Springer, Heidelberg (2009)
17. Zhai, S.: lnvestigation of feel for 6 DOF inputs: isometric and elastic rate control for manipulation in 3D environments. In: Proc. of The Human Factors and Ergonomics Society 37th Annual Meeting, vol. 37(4), pp. 323–327 (1993)
18. Schmitt, B., Raynal, M., Dubois, E., Bach, C.: Extension de la norme ISO 9241-9 au pointage en 3D. In: 23rd French Speaking Conference on Human-Computer Interaction (IHM 2011), pp. 117–120. ACM, New York (2011)
19. Masliah, M.: `http://etclab.mie.utoronto.ca/people/moman/timeandspace/timeandspace.html`
20. Zhai, S., Milgram, P.: Quantifying coordination in multiple DOF movement and its application to evaluating 6 DOF input devices. In: Karat, C.-M., Lund, A., Coutaz, J., Karat, J. (eds.) Conference on Human Factors in Computing Systems (CHI 1998), pp. 320–327. ACM Press/Addison-Wesley Publishing Co., New York (1998)
21. Zhai, S., Buxton, W., Milgram, P.: The "Silk Cursor": Investigating transparency for 3D target acquisition. In: Adelson, B., Dumais, S., Olson, J. (eds.) Conference on Human Factors in Computing Systems (CHI 1994), pp. 459–464. ACM, New York (1994)
22. Schmitt, B., Dubois, E., Raynal, M., Bach, C., Croenne, D.: Evaluation d'une technique d'interaction ubiquitaire pour le pointage de données complexes et spatialisées. In: Journées Francophones Mobilité et Ubiquité (UBIMOB 2011), pp. 1–8 (2011)
23. Schmitt, B., Raynal, M., Dubois, E.: A composite approach to evaluate two interaction techniques for a 3D pointing task. In: IEEE Symposium on 3D User Interfaces, pp. 159–160 (2012)
24. Bangor, A., Kortum, P.-T., Miller, J.-T.: An empirical evaluation of the System Usability Scale. Journal of Human-Computer Interaction, 574–594 (2008)
25. Zhai, S.: User performance in relation to 3D input device design. SIGGRAPH Comput. Graph. 32(4), 50–54 (1998)