

## Head-mounted display versus desktop for 3D navigation in virtual reality: a user study

Beatriz Sousa Santos · Paulo Dias · Angela Pimentel ·  
Jan-Willem Baggerman · Carlos Ferreira ·  
Samuel Silva · Joaquim Madeira

Published online: 28 August 2008  
© Springer Science + Business Media, LLC 2008

**Abstract** Virtual Reality (VR) has been constantly evolving since its early days, and is now a fundamental technology in different application areas. User evaluation is a crucial step in the design and development of VR systems that do respond to users' needs, as well as for identifying applications that indeed gain from the use of such technology. Yet, there is not much work reported concerning usability evaluation and validation of VR systems, when compared with the traditional desktop setup. The paper presents a user study performed, as a first step, for the evaluation of a low-cost VR system using a Head-Mounted Display (HMD). That system was compared to a traditional desktop setup through an experiment that assessed user performance, when carrying out navigation tasks in a game scenario for a short period. The results show that, although users were generally satisfied with the VR system, and found the HMD interaction intuitive and natural, most performed better with the desktop setup.

**Keywords** Virtual reality (VR) · Virtual environments (VE) · Head mounted display (HMD) · User study · Navigation

---

B. Sousa Santos (✉) · P. Dias · S. Silva · J. Madeira  
Departamento de Electrónica, Telecomunicações e Informática, Univ. de Aveiro, Aveiro, Portugal  
e-mail: bss@ua.pt

B. Sousa Santos · P. Dias · S. Silva · J. Madeira  
IEETA—Instituto de Engenharia Electrónica e Telemática de Aveiro, Aveiro, Portugal

A. Pimentel · C. Ferreira  
Departamento de Economia, Gestão e Engenharia Industrial, Univ. de Aveiro, Aveiro, Portugal

J.-W. Baggerman  
Faculty EEMCS, Delft University of Technology, Delft, The Netherlands

C. Ferreira  
CIO—Centro de Investigação Operacional, Universidade de Lisboa, Lisboa, Portugal

## 1 Introduction

Since its infancy in the early 1990s, and in addition to entertainment, Virtual Reality (VR) has found application in different areas, e.g., automotive industry, architecture, medicine and education, in particular for assembly, design review or training [7, 18]. Recent technological advances have allowed the development of lower-cost, lightweight, easy to use systems and VR is entering new areas such as tourism and cultural heritage. Moreover, VR does not necessarily imply immersive stereoscopic visualization: many emerging VR applications are desktop or table-based and not stereoscopic. However, the focus has been placed on achieving the highest possible technical standards, with almost no attention being given to system evaluation involving end-users.

Virtual Reality systems, and the Virtual Environments (VEs) experienced within them, have presented challenges to Human-Computer Interaction (HCI) researchers over many years. The sheer range of different interfaces which might be experienced, and of different behaviors which might be exhibited, makes it difficult to understand the participants' performance within VR/VE and to provide coherent guidance for designers [41]. Moreover, empirical studies on human factors are extraordinarily difficult to plan and carry out, given the large number of variables involved [40].

Navigation is one of the core tasks in VEs [3] but, in general, such environments seem to be poorly designed to support users in this task. There is much interest on how different forms of interaction and various environmental features affect navigational learning [32], such as those found in desktop and immersive setups (such as a Head-Mounted Display—HMD): in the former an abstract interface (e.g., mouse and keyboard) is used, in the latter physical changes of direction are directly mapped into the VE [32].

As asserted by Field [10] for Augmented Reality, and we believe also valid for VR, the development of such applications raises several questions, in addition to simple usability testing: Do users prefer a VR system or an alternative tool? Which tasks and alternative tools should be used for evaluating the usability of a VR application? Which technology advances would likely improve user performance in such systems, is also an open question [34].

In this paper we present the first of a series of user studies we are performing in order to assess the usability of a low-cost VR system, using a HMD, under development at the University of Aveiro, while simultaneously trying to understand if this technology is beneficial to users and in which circumstances.

The importance of navigation in VEs, as well as the evaluation of differences in performance obtained in immersive and non-immersive VEs, lead us to start the usability study by comparing our low-cost VR system with a traditional desktop setup. We assessed the performance and satisfaction of 42 users in both setups, as they performed navigation tasks in the virtual environment while gaming for a short period of time. We also analysed the influence on user performance of several factors, namely, gaming experience, familiarity with 3D environments, and gender.

In the next sections we will describe related relevant work and present our experiment, as well as its main results. Finally we discuss the obtained results and suggest some directions for future work.

## 2 Background

In this section we will first address usability evaluation in VEs, then introduce some fundamental issues related to navigation, and focus on the differences between immersive

and desktop VEs. Finally, we will describe a few user studies which illustrate relevant issues when comparing the usability of the two kinds of systems, and might provide background for anticipating user performance in our study.

## 2.1 Usability evaluation in virtual environments

Since existing usability evaluation methods present limitations for evaluating VEs, there have been attempts to adapt traditional methods for use in VEs and a few efforts to develop structured evaluation methodologies [6, 11, 12].

There are some important differences between evaluation of VE user interfaces and traditional Graphical User Interfaces (GUIs) [3, 6], either concerning physical environment issues—e.g., users wearing a HMD may be standing and cannot see the surrounding physical world, therefore the evaluator must ensure that they will not trip over cables, or bump into objects—, or evaluator related issues—e.g., in VEs touching or talking to the user may cause breaks in the presence, thus if presence is hypothesized to affect performance on the task under evaluation, the evaluator must take care to remain unsensed.

When evaluating user performance, speed and accuracy are easy to measure quantitative values, but other more subjective performance measures should also be considered, such as perceived ease of use, ease of learning, and user comfort. Moreover, since VEs involve user's senses and body in the task, focusing on user-centric performance measures is essential: if a VE does not make good use of human skills, causes fatigue or discomfort, it will not provide overall usability despite its performance in other areas. Thus, VE usability studies must also consider the effects on subjects of simulator sickness and fatigue.

Additional issues are related to the type of evaluation: when performing experiments to quantify and compare the usability of various VE interaction techniques, interface elements, or input and output devices, it is often difficult to know which factors have a potential impact on results. Besides the primary independent variable (e.g., a specific interaction technique), there are a large number of other potential factors that could be included, such as environment, task, system, or user characteristics.

A review of recent VE literature shows that a growing number of researchers and developers are considering usability at some level, and several methods developed for 2D or GUI evaluation have been extended to support VE evaluation, such as controlled experiments, cognitive walkthrough, observation methods, interview, and post-hoc questionnaire. Interesting examples of usability studies of VEs can be found in [1, 8, 14, 16, 21, 24–26, 28, 30, 33, 36] and illustrate the wide range of methods and combinations available.

## 2.2 Navigating in virtual environments

Navigation is one of the core tasks that people perform in VEs: to navigate successfully they must plan their movements using spatial knowledge gained about the environment and stored as a mental map. However, 20% to 30% of the users often have difficulties in navigating in VEs [42]. Navigation in VEs is characterized by a slower acquisition of spatial knowledge than in the real world and, sometimes, extreme difficulty finding places during exploration. It presents challenges such as supporting spatial awareness, providing efficient and comfortable movement between different locations, and making navigation lightweight so that users can focus on more important tasks.

Bowman et al. [3] subdivide navigation into travel (the motor component) and wayfinding (the cognitive component). Travel is a conceptually simple task: the movement of the viewpoint from one location to another. Further, viewpoint orientation is usually

handled in immersive VEs by head tracking, so only techniques for viewpoint positioning need be considered. There are five common metaphors for travel interaction techniques: steering (the most common), physical movement, manual viewpoint manipulation, target-based travel and route-planning, steering being the most common travel metaphor. Besides choosing a travel metaphor, other design issues include velocity control and the use of constraints or guides to aid travel. A taxonomy of the travel task can be found in [2] and [6].

Wayfinding can be described as the cognitive process of defining a path through the environment, thereby using and acquiring spatial knowledge to build up a cognitive map of an environment. The possible lack in VEs of cues present in the real world (e.g., distance and motion), as well as other limitations (e.g., narrow field of view), makes it more difficult for users to develop an understanding of the area [2]. This and the extra degrees of freedom within a VE can easily cause disorientation, thus users should receive wayfinding support, e.g., through large field of view, real-motion, sound, maps, texture, or a horizon [3, 5].

Given the various navigation choices, it might be tempting for designers to include a great number of features and let users choose among them. However, according to an interesting study by Sebok et al. [33], this approach would almost certainly lead to a less usable system than one with a number of reasonable constraints, since effective user-centred design requires achieving a careful balance between providing options and flexibility, while at the same time minimising the user's cognitive burden. Those authors conclude that constraining movement to the  $x$ - $z$  plane and including collision avoidance appear necessary; although it is desirable to limit features offered to the user, additional navigation aids can also be useful.

Due to the impact of navigation issues in the usability of VEs, it has also to be identified how different forms of interaction and a variety of environmental characteristics affect navigational learning [32].

### 2.3 Comparing HMD-based to desktop-based VEs

Virtual environments may be experienced using immersive displays (e.g., HMDs, CAVEs) or desktop displays. In each case the user interfaces have different characteristics and may lead people to perceive VEs in different ways. The HMD was considered the canonical VE display by Bowman et al. [5]; yet, CAVEs and other spatially immersive displays have become more common.

HMDs consist of two LCD screens mounted in a glasses-like device and fixed relative to the wearer's eye position, and portray the virtual world by obtaining the user's head orientation (and position in some cases) from a tracking system. Several HMD features may have an effect on user performance: HMDs may present the same image to both eyes or be stereoscopic, and offer a wide range of resolutions, usually trading off with field of view (FOV). Additional ergonomic issues such as display size and weight, and the ability to adjust various visual parameters are also important.

Concerning interaction, while in desktop VEs people typically use abstract interfaces (e.g., mouse, keyboard, joystick, or a spaceball) to control their translation movements and direction changes, in immersive VEs they also use abstract interfaces to perform translation movements, however physically turn around to change direction.

### 2.4 User studies comparing VR systems using HMDs and desktops

Despite the VR community growing interest in usability related research, direct comparisons of user performance while using HMDs and desktops are not common. We

were able to find six such studies and Table 1 shows some relevant information concerning them: authors, date, evaluation methods, context of evaluation and main findings.

Analysing Table 1, it can be observed that controlled experiments have been used in all studies (using both within-subjects and between-subjects experimental designs [9]), complemented with a questionnaire in two cases, navigation and search being the most studied tasks. For each participant, the duration of the experiment attained several hours in two cases, and four studies included training periods. It is also apparent from Table 1 that most studies were performed in a generic evaluation context (and not applied to a specific situation).

The earliest study analysed was conducted by Pausch et al. [22] and compared a search task between a HMD and a stationary HMD. While it did not involve a desktop, the authors consider their fixed HMD setup contained the essential desktop components: a stationary monitor and a hand-input device. Participants were placed in the centre of a virtual room and instructed to search for a camouflaged target. The study showed that when a target was present, there was no significant performance improvement in the immersive environment. However, when the target was not present, participants in the immersive environment were able to reach that conclusion substantially faster than the participants using the stationary display. A positive transfer of training effect from the immersive environment to the stationary display was also identified, as well as a negative transfer of training effect from the stationary display to the head-tracked environment. Pausch et al. suggest this demonstrates that immersed users had built a better mental model of the environment, being much better on systematic searches, since they remembered where they had already looked in the scene that surrounded them.

A subsequent study by Robertson et al. [29] included an experiment extending the Pausch et al. study to a desktop display, using also a visual search paradigm to examine navigation in desktop VR, both with and without navigational aids. While Pausch et al. had found that non-head tracked users took significantly longer than predicted when the search target was absent, Robertson et al. found that search times matched prediction when the target was absent, indicating that Pausch et al. conclusion does not transfer to desktop systems. According to Robertson et al., this could be due to the fact that the fixed-HMD used by Pausch et al. was quite different from desktop systems: (1) it involved the user wearing a fixed HMD, resulting in users receiving the low resolution of HMDs without their advantage of head-centric camera control; (2) Pausch et al. used an unfamiliar two-handed input device.

In fact, Robertson et al. observed much faster search times than Pausch et al., which might be explained by the desktop display using a smaller visual angle, thus being faster to scan. This suggests a significant advantage for fish-tank and desktop setups over HMD, besides the higher resolution. Finally, and still according to Robertson et al., even if desktop VR is shown to be consistently slower or more demanding than HMD, adding simple navigation aids may eliminate such performance differences.

The study of Ruddle, Payne and Jones was a virtual building walkthrough experiment, designed to compare a HMD with a desktop monitor display [31] in navigation. Participants would learn the layout of large-scale virtual buildings through repeated navigation within two large virtual buildings. A repeated measure design was used, where each participant navigated one building four times using the HMD, and navigated the second building four times using the desktop. There was no significant difference between the two types of display in terms of the distance that participants traveled or the mean accuracy of their direction estimates. However, participants using the HMD navigated the buildings significantly faster and developed a significantly more accurate sense of relative straight-line distance and building layout. The decreased time was attributed to the fact that participants took advantage of the

**Table 1** Studies comparing VR systems using HMDs and desktops (or similar)

Author	Hardware	Context	Methods	Tasks	Main findings
[22]	HMD (6 DOF), fixed HMD	General	Controlled experiment (between-subjects) (24+24=48 students) (training)	Search	Participants using the HMD concluded faster that the searched target is not present. If it is present there was no difference. Positive transfer from HMD to fixed HMD (negative on the opposite direction) was observed
[29]	HMD (3DOF), desktop (extends Pausch's experiment to desktop)	General	Controlled experiments (32; 26; 23 participants) (between-subjects, within subjects)	Search Navigation (with peripheral lenses)	Participants using desktop searched faster when the target was present. No advantage for HMD users was found when the target was absent
[31]	HMD (3DOF, non-stereo), desktop	General	Controlled experiment, (within-subjects) (12 students) (hours) (training <1 h) Questionnaire	Navigation (in large-scale VEs)	Participants navigated quicker with the HMD, they spent less time stationary, looking around more while moving
[20]	HMD (3DOF, stereo), desktop/joystick	General	Controlled Experiments (within-subjects) (36 subjects)	Physically replicate a virtual sculpture	No difference between the desktop/joystick and HMD was found
[32]	HMD (3DOF, mono), desktop (mono)	General	Controlled experiment (within-subjects) (12 participants) (training) (5–6 h)	Navigation (proprioceptive feedback and environment characteristics) Search	No differences in the rate at which knowledge developed with the two displays were observed. Straight line distances were estimated more accurately with the desktop
[27]	HMD (6DOF), fish tank, fish tank with haptics	Applied	Controlled experiment (between-subjects) (14+13+13=30 participants) (training) Questionnaire	Generic Volume Visualization tasks (identification, judgement of size, shape and connectivity)	Participants using fish tank (in both conditions) were faster and more accurate than the HMD group

natural, head-tracked interface provided by the HMD in ways that included “looking around” more often while traveling through the VEs, and spending less time stationary while choosing a direction in which to travel.

Mizell et al. [20] aimed at determining whether immersive VR technology gives users a measurable advantage over more conventional display methods, when visualizing complex 3D geometry. Participants were shown an abstract rod sculpture in a variety of display and display-control modes, were tasked with assembling a physical replica of the sculpture they were visualizing, and were scored on the speed and accuracy with which they assembled it. No statistically significant difference between the desktop/joystick and the HMD was found; however, the authors consider these results could be confounded by the idiosyncrasies of the equipment.

Ruddle and Péruch [32] investigated the effect of proprioceptive information and environmental characteristics on spatial learning, when participants repeatedly navigate complex 3D virtual mazes. Proprioceptive information, provided by viewing the mazes using a HMD, was found to have little effect. This is in contrast to the study by Ruddle, Payne and Jones [31], and no satisfactory explanation was offered. According to the authors, a possible limiting factor to the benefit of proprioceptive information was the lower spatial resolution of the HMD, which might have hindered the estimation of distance. The primary environmental characteristics in this study were layout orthogonality (using paths that intersected at either oblique or 90° angles), lines of sight (controlled using computer-generated “fog”), a visually defined perimeter and global landmarks. The results indicated that both orthogonality and lines of sight are similarly important. Global landmarks promoted a similar rate of spatial learning to a visual perimeter.

Using the results of an empirical study, Qi et al. [27] developed guidelines for the choice of display environment for four specific, but common, volume visualization problems: identification and judgment of the size, shape, density and connectivity of objects present in a volume. User performance in three different stereo VR systems was compared: (1) HMD; (2) fish tank; and (3) fish tank augmented with a haptic device. HMD participants were placed “inside” the volume and walked within it to explore its structure. Fish tank and haptic participants saw the entire volume on-screen and rotated it to view it from different perspectives. Response time and accuracy were used to measure performance. Results showed that the fish tank and haptic groups were significantly more accurate at judging the shape, density, and connectivity of objects, and completed the tasks significantly faster than the HMD group. As guidelines, the authors recommend an “outside-in” system for the analyzed tasks, since it offers both overview and context, two visual properties that are important for the volume visualization tasks, as well as haptic feedback, if possible, since it aided comprehension, according to the user’s opinion.

In a nutshell, the results obtained from the few user studies we were able to find comparing user performance in immersive VEs, using HMDs and desktop VEs, did not provide much help to anticipate how users would perform in the two conditions we intended to compare.

### 3 The virtual reality system and virtual environment

Our low-cost VR system consists of stereo HMD *i-glasses SVGA Pro* with a resolution of 800×600 pixels, stereoscopic capabilities, 26° of field of view and a frame rate of 60 Hz or 120 Hz (corresponding to mono or stereo), an orientation sensor (tracker) *InterTrax 2* from *InterSense* with three degrees of freedom (yaw, pitch and roll) and a PC with a Nvidia



GeForce FX5950 graphics card. The desktop setup we used had a 17" monitor with a resolution of 800×600 pixels.

A number of possible VEs were considered as potentially suitable to study the usability of our system. We chose a maze, since it would be simple to construct, using existing tools, and provide a good way to test some forms of navigation. On the other hand, a game scenario could be made to incorporate several navigation tasks that users would have to perform spontaneously, while instilling a sense of competition between users, as observed by Griffiths et al. [13]. Thus, we finally decided to develop a game based on a maze having very similar corridors, in which users had to navigate in order to find objects (see Fig. 2). There were 21 objects to be located, which were floating at eye level, and users had only to collide with them. Audio feedback was given whenever users caught an object and the attained score was shown.

While immersed in the VE and playing the game, users were compelled to perform a series of tasks we had defined as important, although in different order and a different number of times among participants. We believe that, since it was not necessary to ask users to perform particular tasks, it made the experiment simpler and more "natural". Also, no instructions had to be given to the users during the game, which is important in order not to disrupt their involvement and sense of presence.

There are two fundamental ways to navigate around VEs: constrained by gravity (e.g., positioned on the ground) or "flying". We chose the former, since it is more "natural" and usually results in less user disorientation [37]. Navigation tests while grounded should contain simple actions such as moving forwards in a straight line and navigating around corners, as well as more complex manoeuvres, for instance walking through doorways (which requires correct judgement of distances between viewpoint and the doorway) [13]. Thus, we designed our VE as to compel users to perform simple and complex navigation tasks (e.g., forward motion, navigating around corners and through doorways).

As mentioned before, navigation can be subdivided into the travel and wayfinding components. We used steering (the continuous specification of direction of motion) as travel metaphor, handled by head tracking when using the HMD and through the mouse on the desktop setup.

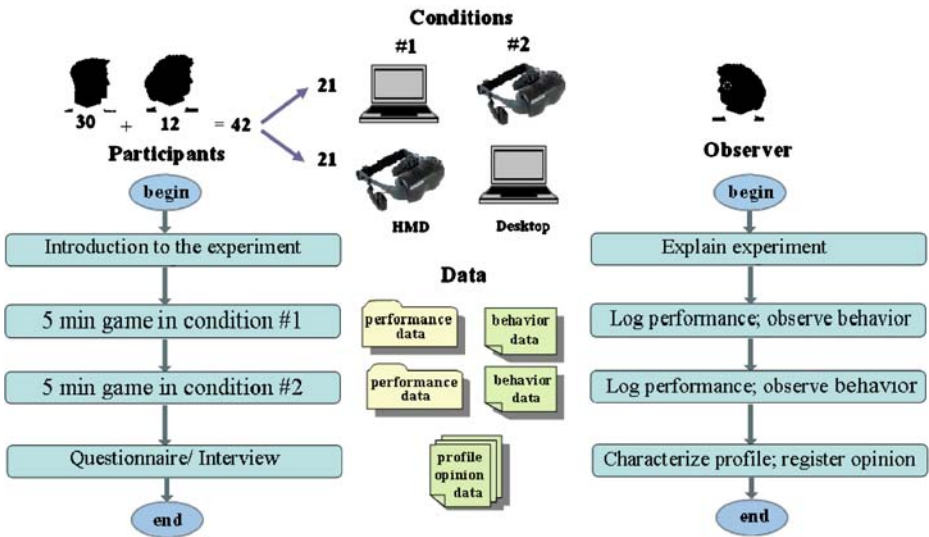
Despite the possible positive impact on the users' performance of navigation aids (e.g., maps, landmarks or trails) and the recommendations of some authors, we decided not to offer any of those aids; however the corridors of the maze are orthogonal to each other, which should make its traversal easier [32].

The complex structure of the maze, the fact that corridors are very similar to each other, and the absence of navigation aids, makes it virtually impossible to build a mental map of the environment. We anticipated that most users would just navigate randomly through the maze, and thus their performance would be mostly due to travel and not to wayfinding.

#### 4 The experiment

Figure 1 shows the observer and participants work flow along the experiment, as well as the hardware used in each experimental condition and collected data. As mentioned, the game developed for this experiment was played in two different modes. In immersive mode (using a HMD) a user's head movements and the mouse buttons were used to control the walking direction (right button forward, left button backward), each user being allowed to hold the mouse in his/her dominant hand (see Fig. 2). The desktop version was played in a way similar to most point and shoot games: the mouse was used to select the moving direction, and two keyboard keys were used to go forward or backward.





**Fig. 1** The experiment: observer and participants work flow, hardware and data collected along the experiment

#### 4.1 Participants

Since we were not conducting our usability study in the context of a specific application, the target user population was not well defined, and the best course of action was to evaluate the most diverse user population possible in terms of age, gender, physical characteristics, etc. Therefore, we asked for the collaboration of 42 volunteer users (12 women and 30 men), aged from 14 to 40 (the majority between 16 and 30). Many were graduate or post-graduate students in our department, but there were also lecturers, administrative staff and high-school students. None of the users had experience in using VR systems. Some of them had gaming experience.



**Fig. 2** User playing in VR mode and aspect of the virtual environment

Every user played the game in both setups (using the HMD and desktop). The experiment started with a short introduction to the game, its goals, the hardware components (HMD basically) used, and a reference to the questionnaire to be completed at the end of the experiment. Users were also informed that they could abandon the experiment at any time. After this initial presentation, each user started playing the game without any previous training. During the game, an observer was monitoring the user performance and taking down relevant information.

#### 4.2 Hypothesis and variables

Applying to VEs what is known about navigational learning in the real world, we would expect that HMD usage (implying physical rotation) would bring greater navigational benefits than using an abstract interface (desktop), and experimental studies using a variety of systems show that this predicted difference does occur in simple navigational tasks. However, contradictory results have been found when participants performed a more complex navigational task and searched for objects in a large scale space [31].

Based on the literature review carried out, we were not able to anticipate what would be the result of our study and therefore we started from the general simple hypothesis (H0) that the performance and satisfaction of users would be similar in both conditions: using the HMD and using the desktop (from now on named VR and VD, respectively).

Once established the hypothesis, the independent and dependent variables are in general easily defined [9]. In our study we used as independent (or input) variable the type of system (with two levels: VR and VD) and as dependent (or output) variables, user performance related variables.

Since experience with video and computer gaming, as well as gender, have been studied as possibly influencing users' performance in navigation [39], we considered them as secondary variables and examined their influence. We also deemed experience in viewing 3D computer scenes as potentially influencing user performance and, thus, used it as a secondary variable as well. In addition, we investigated if experience with the game in one condition (VR or VD) increased user performance on the other condition, to try to ascertain whether there was knowledge transfer between conditions.

#### 4.3 Tasks

During the game, users had to navigate in the virtual maze for a period of 5 min, to locate and collect as many objects as possible (see Fig. 2). Such limited duration was selected to minimize participants' discomfort, yet allowing them to perform a reasonable number of tasks.

As mentioned before, the VE was designed to compel users to spontaneously perform a set of navigation tasks while playing the game:

- forward motion
- backward motion
- cornering
- navigating onto a specific point
- navigating through doorways.

While the first three are simple navigation tasks, the last two are more complex and require the correct judgement of distances to destination. Users had only to perform a very simple object interaction task to catch objects, listening to audio feedback and increasing their score when each task was successfully completed.

The game was previously tested by the authors, as well as by several other people with different levels of experience with the VR system, in order to detect possible issues that might make the game too easy or too difficult for the users, such as speed control, object visibility, doorframe size and corridor width. A few adjustments were done afterwards.

#### 4.4 Performance measures and other collected data

Even the simpler tasks may be difficult for a new VE participant. Observation has shown that users frequently collide with doorframes or hatches, leading to frustration and annoyance, which can have an impact on user performance and satisfaction. Taking this into consideration, user performance was assessed via a set of quantitative measures automatically logged by the system during the game: number of collected objects, number of collisions with walls, walked distance, average speed, and total gaming time. Besides them, additional relevant information was recorded by an observer, concerning users' behavior, difficulties found and performance during the game (e.g., the number of times each user failed to “catch” an object in a first attempt). The observer could always see the image shown to the user, and tried to be as inconspicuous as possible, in order to not disturb or distract the user.

After playing in both conditions (VR and VD), users were presented a questionnaire with a few questions about their profile (as age, gender, profession, game playing habits, experience in using 3D), as well as about their satisfaction, opinion and preferences regarding the two playing modes. Moreover, we collected comfort related data following the recommendations in [4]. After answering the questionnaire, users were interviewed, and some additional questions and suggestions were asked.

#### 4.5 Experimental design

As mentioned, we started from a general simple hypothesis that the performance and satisfaction of users were similar in both environments and used a within-groups experimental design [9], i.e., all participants performed under both conditions, VR and VD. This experimental design has the advantages of requiring fewer participants, as well as reducing the effects of individual differences. To avoid a possible bias on the results due to learning or boredom, the order in which the conditions were tackled was varied among users, by randomly dividing them in two groups: the first started by the VR and the second by the VD experiments.

## 5 Results

We present now the main results from the analysis of the data collected during the experiment, obtained through Exploratory Data Analysis, non-parametric tests and Multivariate Analysis, all performed using STATISTICA [35].

The first variables analysed were the number of caught objects, number of collisions, distance and speed. After a preliminary Exploratory Data Analysis (EDA) using box-plots [17], we decided to test the equality of medians by means of non-parametric tests, since none of the data had normal distribution. Figure 3 shows, on the left, the box-plots corresponding to the number of objects caught by all participants and, on the right, the number of collisions. The median values of the number of caught objects were 14.5 in the immersive environment (VR in the boxplots) and 17 for the desktop (VD in the box-plots). A Wilcoxon Matched Pairs Test [15] rejected the hypothesis of equality of medians ( $p=0.000002$ ). This

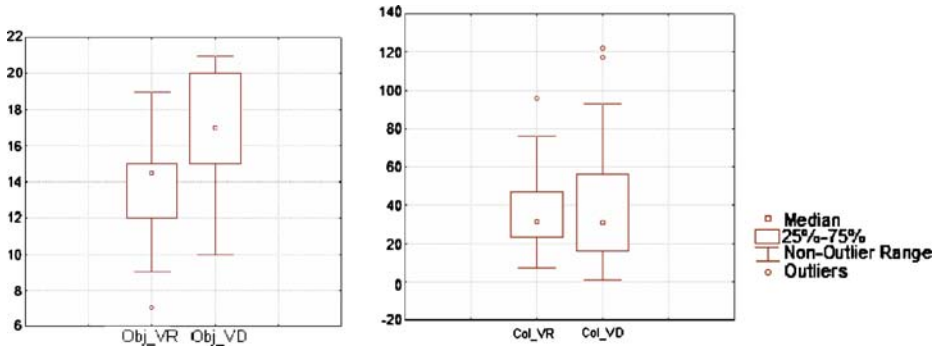


Fig. 3 Number of objects caught by users (left), and number of collisions (right)

means that the difference between the number of objects caught by the users in both environments is statistically significant and cannot be due to chance. On the contrary, the median values of the number of collisions (31.5 for VR and 31 for VD), are not significantly different, since a test of the same type didn't reject the hypothesis of median equality ( $p=0.880730$ ).

The box-plots corresponding to distance and average speed are shown in Fig. 4. Also in these two cases, Wilcoxon Tests rejected the hypothesis of equality of median values of distance (543.5 for VR and 819.5 units for VD;  $p<0.000000$ ) and speed (6.6 units for VR and 9.9 for VD;  $p<0.000000$ ).

From these results we can conclude that users performed globally better when using the desktop (VD) than when using the immersive environment (VR), i.e., they caught more objects, attained higher average speeds and walked larger distances, with roughly the same amount of collisions.

In spite of the fact that the main independent variable of this experiment was the test environment (with two different levels: VR and VD), some secondary variables that could have been influencing the obtained results were identified from the onset of the experiment, as mentioned before; thus, we decided to investigate their influence on the obtained results. These variables are related to the users' profile (as gender, experience in computer games, and viewing images of 3D data, models or scenes), or to working conditions (as learning effects from one environment to the other). Correlation between walked distance and average speed in both environments is very high (1 for VR and 0.96 for VD) as expected, since the great majority of users played for the same amount of time (5 min), as a

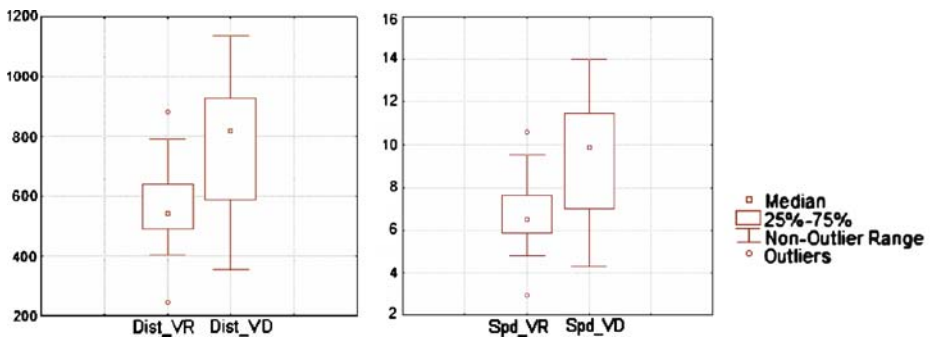


Fig. 4 Distance walked by users (left), and average speed (right)

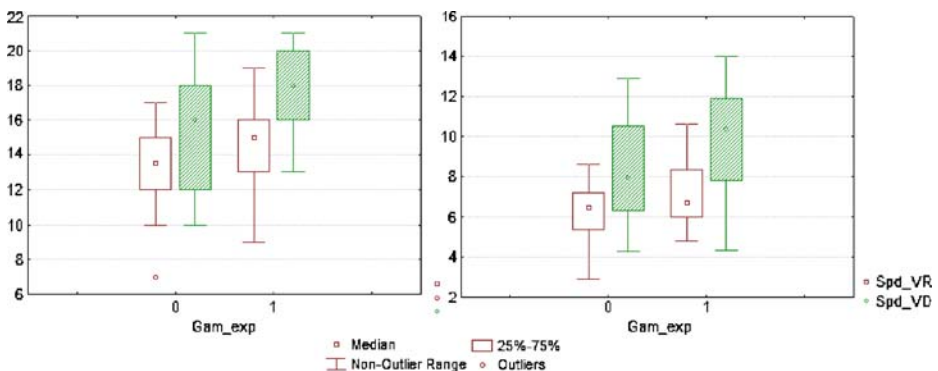
consequence we considered only one of these variables. Moreover, as the number of collisions is not significantly different in both environments, we studied the number of objects caught and the average speed categorized by gaming experience (Fig. 5), experience with 3D (Fig. 6), and gender (Fig. 7).

In order to study the influence of those variables, users were categorized according to their experience, taking into account their answers to the questionnaire and making groups of users as balanced as possible. Thus, in spite of having asked users how they classified their experience in computer games and 3D (not experienced, moderately experienced, and highly experienced), we have categorized experience in computer games using only two levels (not experienced—0, experienced—1), and with 3D using the three levels (not—0, moderately—1 and highly experienced—2). This categorization produced groups having a number of users varying from 11 to 18 which, for the statistical analysis purposes, seems reasonable.

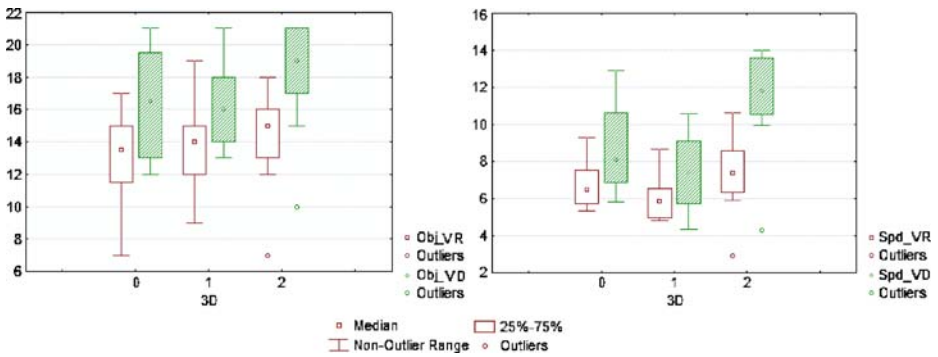
Figure 5 shows the number of caught objects and the average speed for the two user groups according to their experience with computer games. We can observe that both variables have higher values for VD (hatched boxes) than for VR, for all user groups, as it was already found for all users in Figs. 3 and 4. Analysing the results obtained by users with and without experience, we notice that, for VR, neither the number of objects caught nor the average speed vary much between these groups. However, the same is not true for the desktop (VD), in which case users having experience caught significantly more objects ( $p=0.01$ ) and attained higher speeds ( $p=0.04$ ), than non-experienced users.

Figure 6 shows the number of caught objects and the average speed for the three user groups according to their experience with 3D. As in the previous case, both variables have higher values for VD than for VR, for all user groups, as it was already found for all users in Figs. 3 and 4. The number of objects caught by the users in both environments (VR and VD) was not influenced by this categorization. On the other hand, the equality of average speed among the three groups was rejected (Kruskal–Wallis:  $p=0.03$  for VR and  $p=0.004$  for VD). These results could mean that previous experience with 3D was relevant for the speed attained in both environments, but experience with computer games was relevant only when using the VD environment.

Regarding gender, Fig. 7 shows the number of objects caught and average speed attained by men and women. The box-plots suggested, and it was statistically supported (by a Mann–Whitney test), that both groups seem to have similar performances in the VR environment; however, men outperformed women in the VD. A note of caution is due to the fact that the two groups are not balanced (12 women and 30 men) and men have more



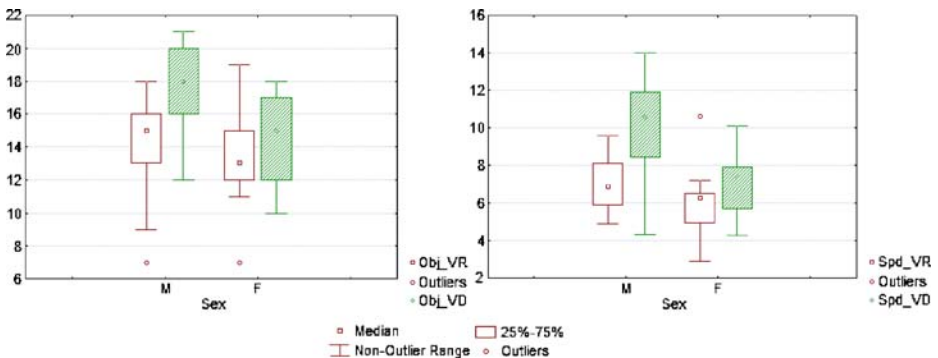
**Fig. 5** Number of objects caught by users (*left*), and average speed (*right*), categorized by experience in computer games



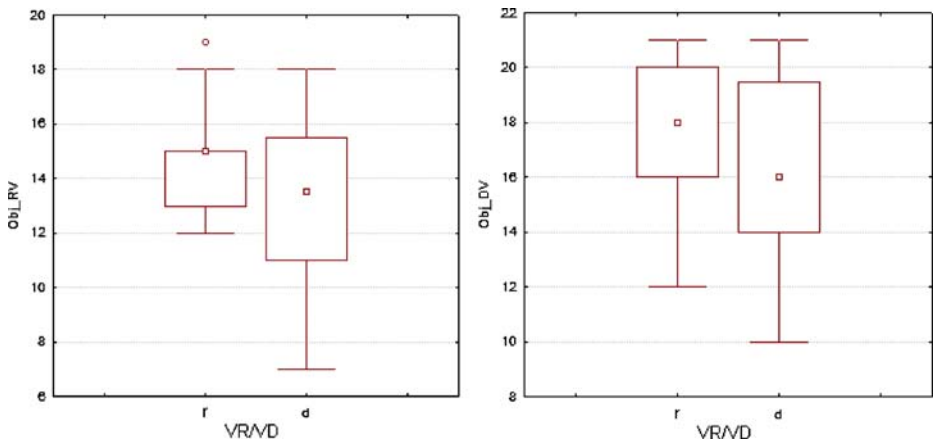
**Fig. 6** Number of objects caught by users (*left*), and average speed (*right*), categorized by experience in using 3D

experience in computer games, which can make invalid the apparent result that women in general could perform better in the VR environment than in the VD.

In order to alleviate learning effects between the two environments, half of the users started the experiment by the VD condition and the others started by the VR condition. Nevertheless, we tried to investigate if there was an increase on the attained speed, or on the number of objects caught, on the second time users were playing the game, comparing their performances with the performances of users playing the game on the same environment but at a first time. This comparison was done analysing the box-plots of Figs. 8 and 9, and performing Mann–Whitney tests in order to assess if the differences between medians, observed in the box-plots, are statistically significant. These figures show the number of objects caught (Fig. 8) and the average speed attained (Fig. 9), by the users in each environment, categorized by the environment where they started the experiment (i.e., in which they have played the game for the first time). Observing Fig. 8 we notice that users who played first in the VR environment caught slightly more objects in both environments than users starting by the VD environment, however this was not statistically significant, thus suggesting that there was no knowledge transfer concerning the location of objects in the maze. This was expected, since the environment had been designed to make the acquisition of a mental map difficult. On the other hand, the average speed was generally lower in the environment users experienced first, as can be observed on Fig. 9, which means that users moved faster as they got more experienced.



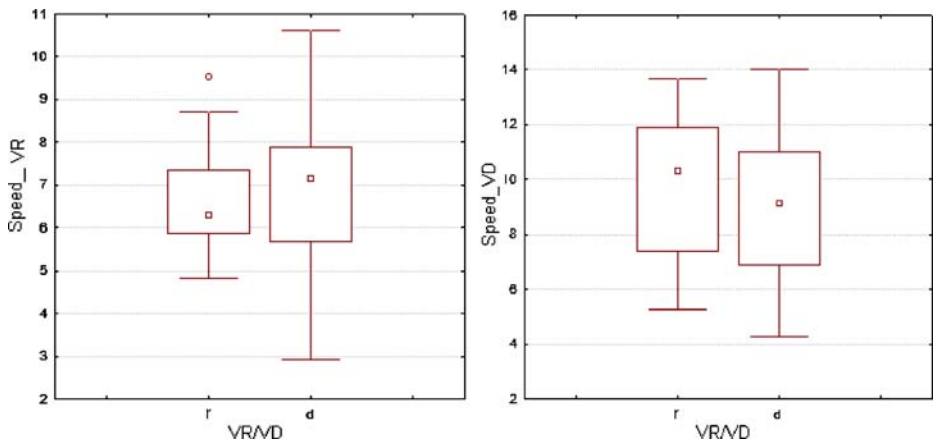
**Fig. 7** Number of objects caught by users (*left*), and average speed (*right*), categorized by gender



**Fig. 8** Number of objects caught by users in the VR (*left*) and in the VD (*right*) environments, categorized by the environment that was used first (*r* immersive environment, *d* desktop)

After the experiment users were asked about some other relevant issues: whether they had felt disoriented or sick, and their degree of satisfaction with the experience in both environments. The results were that more users felt disoriented and sick in the VR environment (19 and 25 users respectively) as compared to the desktop (five and ten users); however, the degree of satisfaction was comparable (31 users enjoyed both experiences very much).

Finally, in informal conversations afterwards, many users mentioned that the cables connecting the hardware they were wearing in the VR environment (glasses and mouse) were uncomfortable and might have contributed to their disorientation, hence suggesting the use of wireless devices. Some users reinforced this suggestion explaining that they had felt embarrassed to play the game using the VR environment in public, mainly because the cables were often entangled around them as they turned while exploring the maze.



**Fig. 9** Average speed attained by users in the VR (*left*) and in the VD (*right*) environments, categorized by the environment that was used first (*r* immersive environment, *d* desktop)



## 6 Discussion

Contrary to other studies (e.g., [28]), we did not intend to investigate the benefits of immersion, but merely how users would perform when using our low-cost VR system (using a HMD), as compared to a simple desktop setup. In our experiment global user performance was better for the desktop setup, which might result from the fact that participants were using a standard configuration for game interaction known to many of them, whilst most of them had never used a HMD before. However, some performance results were not very different (for instance, in terms of median, the main goal, catching objects, just differs from 14.5 to 17). Moreover, a few users, mainly those who seldom play computer games, even performed better with the HMD (caught more objects).

Williams et al. [39] found gamers and non-gamers did not significantly differ on errors in their experiment, meaning that skills obtained with game practice did not transfer, from the joystick/console interfaces and small screen virtual environments typically provided by the games, to the immersive HMD/walking system they had used. It is possible that a similar phenomenon has happened in our case, since we noticed that neither the number of objects caught, nor the speed attained vary much between the user groups with and without gaming experience using the HMD; yet, the same is not true for the desktop, in which case experienced users caught significantly more objects and attained higher speeds than non-experienced users.

Concerning possible learning effects, our results suggest that there is a learning effect on the attained speed: users playing the game for the second time attained higher average speeds. Moreover, this effect seems to exist in both directions (from the desktop to the immersive environments and vice-versa). If, on one hand, this seems to contradict the idea presented in [22], that users only transfer experience from the HMD to the fixed HMD environment, it agrees with the findings of [29], which used conditions more similar to the ones used in our study.

Our findings confirm some of the results of a set of experiments described in [19], namely that, for certain applications it may be more appropriate to use desktop systems (particularly if VR induced sickness is a primary concern), and participants enjoyed using HMDs.

As in some other user studies involving navigation in VEs (e.g., [38]), some gender differences were observed; however, in our case this could be due to two different reasons: the fact that the number of women and men was not balanced, or the fact that the users had to perform navigation tasks in a complex, self similar, VE without any landmarks, which seems to be adverse to the strategy usually followed by women [23].

The characteristics of the maze make it virtually impossible to build a mental map of the environment; we assume most users just navigated randomly through the maze, and probably their performance was mostly due to travel and not to wayfinding. Nonetheless, it would be interesting to verify this idea by analysing the paths followed by users as they traverse the maze.

Even if we believe our results are statistically significant, our experiment is an abstract user study and, thus, it is not obvious how the results transfer to a real application domain. Another limitation of this study is the fact that the results obtained from categorization (according to several secondary variables) might be less relevant than others, given that some of the groups have a relatively small number of subjects.

After observing and interviewing our users, we believe that supporting technologies for Immersive Virtual Reality still face great usability challenges. Image brightness and contrast are still better in a desktop display, which also induces less disorientation and sickness as compared to HMDs; in addition, hardware cables also increase discomfort. In fact, cables were the most unsatisfactory part of the VR system: many users commented on

this issue as being a major difficulty, which confirms the observation of [37] concerning the VR system they used in a user study.

In summary, the short duration of the game (5 min), the lack of a training period, the low FOV of the HMD, as well as the cables, all may have hindered users to experience a possible positive effect on the spatial knowledge obtained from the proprioceptive feedback, while using the HMD.

## 7 Conclusions and future work

We compared, through a controlled experiment, the performance of 42 users navigating in a complex virtual environment (a maze) using HMD and desktop setups. The large number of users (compared to most similar studies found in literature), as well as the experimental design and the statistical methods used, give us reasonable confidence that our findings are statistically significant.

The main conclusion of our experiment is that global user performance was better for the desktop setup. Nevertheless, some performance results were not very different between both setups, and those users who seldom play computer games generally performed better with the HMD.

Another important result is that users in general enjoyed using the HMD Virtual Environment, yet most of them commented that cables were a major difficulty and were the most unsatisfactory part of the VR system.

An issue that would need further investigation is whether users would improve their performance as they gain experience with the HMD setup. Thus, repeating this experiment by letting users familiarize themselves with this setup and then gaming for longer periods would, perhaps, clarify if the use of a HMD would have more benefits over the desktop in such circumstances.

As future lines of research, and since navigation is a core task in many important VR applications, it would be interesting to extend this experiment in order to study orientation tasks and potential ways of increasing the performance of users in a low-cost immersive Virtual Reality system as ours. The use of auralization seems likely to help users to find their way in a complex environment (as the maze used) and, since our system allows the use of 3D sound, it would be interesting to investigate if, and in which conditions, its use could help users navigate in the VE. Another interesting future work could be to study the impact of various navigation aids (as maps and trails) on user performance.

Finally, minimizing the problems related to the cables by making our system as wireless as possible, would be an important enhancement.

**Acknowledgements** The authors wish to thank Pauline van der Horst, Frank van Huussen and Dr. Rafael Bidarra for their suggestions, as well as all the participants in the study.

## References

1. Ardito C, Constabile MC, De Angeli A, Pittarello F (2007) Navigation help in 3D worlds: some empirical evidences on use of sound. *Multimedia Tools Appl* 33:201–216 doi:10.1007/s11042-006-0060-0
2. Bowman D, Davis E, Hodges L, Badre A (1999) Maintaining spatial orientation during travel in an immersive virtual environment. *Presence Teleoperators VR* 8(10):618–631 doi:10.1162/105474699566521

3. Bowman D, Kruijff E, LaViola J Jr, Poupyrev I (2001) An introduction to 3D user interfaces design. *Presence Teleoperators VR* 10(1):96–108 doi:10.1162/105474601750182342
4. Bowman D, Gabbard J, Hix D (2001) Usability evaluation in virtual environments: classification and comparison of methods. Technical Report TR-01–17, Computer Science, Virginia Tech
5. Bowman D, Datey AA, Ryu, YS, Farooq U, Vasnaik O (2002) Empirical comparison of human behavior and performance with different display devices for virtual environments. In: Proceedings of human factors and ergonomics society annual meeting, Baltimore, pp 2134–2138
6. Bowman D, Kruijff E, LaViola J Jr, Poupyrev I (2005) 3D user interfaces: theory and practice. Addison Wesley, Los Angeles
7. Brooks F (1999) What's real about virtual reality. *IEEE Comput Graph Appl* 19(6):16–27 doi:10.1109/38.799723
8. Demiralp C, Jackson C, Karelitz D, Zhang S, Laidlaw D (2006) CAVE and fish tank virtual-reality displays: a qualitative and quantitative comparison. *IEEE Trans Vis Comput Graph* 12(3):323–330 doi:10.1109/TVCG.2006.42
9. Dix A, Finley J, Abowd G, Russell B (2004) Human computer interaction, 3rd edn. Prentice Hall, Englewood Cliffs, NJ
10. Field M (2004) Usability and collaborative aspects of augmented reality. *Interaction* 11(6):11–15 doi:10.1145/1029036.1029044
11. Gabbard J (1998) A taxonomy of usability characteristics in virtual environments. MSc thesis, Virginia Polytechnic Institute and State University
12. Gabbard J, Hix D, Swann JE II (1999) User-centred design and evaluation of virtual environments. *IEEE Comput Graph Appl* 19(6):51–59 doi:10.1109/38.799740
13. Griffiths G, Sharples S, Wilson J (2006) Performance of new participants in virtual environments: the Nottingham tool for assessment of interaction in virtual environments (NAÏVE). *Int J Human Comput Stud* 64(3):240–250
14. Gruchalla K (2004) Immersive well-path editing: investigating the added value of immersion. *Proc IEEE Virtual Real 1999*:157–164
15. Hettmansperger T, McKean J (1998) Robust nonparametric statistical methods. Kendall's library of statistics, vol 5. Arnold, London
16. Hix D, Swan E, Gabbard J, McGee M, Durbin J, King T (1999) User centered design and evaluation of a real-time battlefield visualization virtual environment. *Proc IEEE Virtual Real 1999*:96–103
17. Hoaglin D, Mosteller F, Tukey J (1983) Understanding robust and exploratory data analysis. Wiley, New York
18. Israel J, Naumann A (2007) Human-machine interaction in virtual environments—recent developments and industrial applications. *MMI Interakt* 12(April):1–2
19. Karaseitanidis J, Amditis A, Patel H, Sharples S, Bekiaris E, Bullinger A et al (2006) Evaluation of virtual reality products and applications from individual, organizational and societal perspectives—the “VIEW” case study. *Int J Human Comput Stud* 64:251–266
20. Mizell D, Jones S, Slater M, Spanlang B (2002) Comparing immersive virtual reality with other display modes for visualizing complex 3D geometry. <http://www.cs.ucl.ac.uk/research/vr/Projects/Immersion/Experiment2/paper.pdf> (online September, 2007)
21. Narayan M, Waugh L, Zhang X, Baina P, Bownan D (2005) Quantifying the benefits of immersion for collaboration in virtual environments. *Symp Virtual Real Softw Technol* 05:78–81 doi:10.1145/1101616.1101632
22. Pausch R, Proffitt D, Williams G (1997) Quantifying immersion in virtual reality. In: Proceedings of 24th Annual Conference on Computer Graphics and Interactive Techniques, pp 13–18
23. Parush A, Berman D (2004) Navigation and orientation in 3D user interfaces: the impact of navigation aids and landmarks. *Int J Human Comput Stud* 61:375–395 doi:10.1016/j.ijhcs.2003.12.018
24. Patel H, Stefani O, Sharples S, Hoffmann H, Karaseitanidis I, Amditis A (2006) Human centred design of 3D interaction devices to control virtual environments. *Int J Human Comput Stud* 64:207–220
25. Polys N, Seocho K, Bowman D (2005) Effects of information layout, screen size, and field of view on user performance in information-rich virtual environments. In: Proceedings of the ACM Symposium on Virtual Reality Software and Technology'05, pp 46–55
26. Prabhat M, Forsberg A, Slater M, Wharton K, Katzourin M (2008) A comparative study of desktop, fish tank and cave systems for the exploration of volume rendered confocal data sets. *IEEE Trans Vis Comput Graph* 14(3):551–563 doi:10.1109/TVCG.2007.70433
27. Qi W, Taylor R, Healey C, Martens JB (2006) A comparison of immersive HMD, fish tank VR and fish tank with haptics displays for volume visualization. In: Proceedings of the 3rd Symposium on Applied Perception in Graphics and Visualization, pp 51–58

28. Raja D, Bowman D, Lucas J, North C (2004) Exploring the benefits of immersion in abstract information visualization. In: Proceedings of Immersive Projection Technology Workshop, <http://infovis.cs.vt.edu/papers/IPT2004-Dheva.pdf> (online August 2007)
29. Robertson G, Czerninski M, van Dantzich M (1997) Immersion in desktop virtual reality. In: Proceedings of the 10th Annual Symposium on User Interface Software and Technology, ACM UIST'97, pp 11–19
30. Robinson G, Ritchie J, Day P, Dewar W (2007) System design and user evaluation of co-star: an immersive stereoscopic system for cable harness design. *Comput Aided Des* 39:245–257 doi:10.1016/j.cad.2006.12.001
31. Ruddle R, Payne S, Jones D (1999) Navigating large-scale virtual environments: what differences occur between helmet-mounted and desk-top displays? *Presence Teleoperators VR* 8(2):157–168 doi:10.1162/105474699566143
32. Ruddle R, Péruch P (2004) Effects of proprioceptive feedback and environmental characteristics on special learning in virtual environments. *Int J Human Comput Stud* 60:299–326 doi:10.1016/j.ijhcs.2003.10.001
33. Sebok A, Nystad E, Helgar S (2004) Navigation in desktop virtual environments: an evaluation and recommendations for supporting usability. *Virtual Real* 8:26–40
34. Smith R, Fisher B, Sandin D, Westin S (2005) VR—a reality check? *Proc IEEE Conf Virtual Real* 2005:301
35. STATISTICA 6.0, <http://www.statsoft.com>, (on-line September/2007)
36. Steed A, Parker C (2005) Evaluating effectiveness of interaction techniques across immersive virtual environmental systems. *Presence Teleoperators Virtual Environ* 14(5):511–527 doi:10.1162/105474605774918750
37. Usoh M, Arthur K, Whitto M, Bastos R, Steed A, Slater M et al (1999) Walking > walking-in-place>flying, in virtual environments. *Proc SIGGRAPH* 99:359–364
38. Waller D, Hunt E, Knapp D (1998) The transfer of spatial knowledge in virtual environment training presence. *Teleoperators Virtual Environ* 7(2):129–143 doi:10.1162/105474698565631
39. Williams B, Narasimham G, McNamara T, Carr T, Rieser J, Bodenheimer B (2006) Updating orientation in large virtual environments using scaled translational gain. In: Proceedings of the 3rd Symposium on Applied Perception in Graphics and Visualization, APCV2006, pp 21–28
40. Wilson J (2006) Interaction with virtual environments. *Int J Human Comput Stud* 64:157
41. Wilson J, D'Cruz M (2006) Virtual and interactive environments for work of the future. *Int J Human Comput Stud* 64:158–169
42. Witmer B, Bailey J, Knerr B, Parsons K (1996) Virtual spaces and real world places: transfer of route knowledge. *Int J Hum Comput Stud* 45:413–428 doi:10.1006/ijhc.1996.0060



**Beatriz Sousa Santos** is Associate professor at the Department of Electronics, Telecommunications and Informatics, University of Aveiro, Portugal. She graduated and received a PhD in Electrical Engineering in 1989, both from the University of Aveiro. Her main research interests are in the areas of Data and Information Visualization and Virtual Reality.



**Paulo Dias** is Assistant professor at the Department of Electronics, Telecommunications and Informatics, University of Aveiro, Portugal. He graduated and received a PhD in Electrical Engineering in 2003, both from the University of Aveiro. His main research interests are in 3D Reconstruction, Computer Vision, Visualization and Virtual Reality Systems.



**Angela Pimentel** received her MSc in Information Management in 2007 from the University of Aveiro. She participated in this work in the scope of her thesis and her main research interests are Virtual Reality and Usability.



**Jan-Willem Baggerman** was a Computer Science student at the Faculty EEMCS, Delft University of Technology, Delft, The Netherlands and visited the University of Aveiro as an Erasmus free-mover where he participated in this work.



**Carlos Ferreira** graduated in Electrical Engineering, received a MSc degree in Statistics and Operational Research from the University of Lisbon and a PhD in Mathematics from the University of Aveiro in 1998. He is currently an Associate professor with the Department of Economics, Management and Industrial Engineering at the University of Aveiro, Portugal, where he is director of the Information Management MSc. His teaching and research interests are in Operations Research, Data Analysis and Information Management.



**Samuel Silva** is a PhD student at the University of Aveiro, Portugal, where he also received his degree in Electrical Engineering in 2003 and a MSc in Electrical Engineering in 2007. Since 2003 he has been working in the Institute of Electronics and Telematics Engineering of Aveiro (IEETA). His research interests include geometric modeling, quality assessment of polygonal meshes, computer graphics, data visualization and medical imaging of the heart.



**Joaquim Madeira** is Assistant professor at the Department of Electronics, Telecommunications and Informatics of the University of Aveiro, Portugal. He graduated in Electrical Engineering in 1986 and received a MSc degree in Computer Science in 1991, both from the University of Coimbra, Portugal. In 1998 he received his PhD degree in Computer Science from the Darmstadt University of Technology, Germany. Geometric Modeling is his main research area, with focus in modeling and visualization using polygonal meshes, in particular for medical applications.