



# Usability in virtual reality: evaluating user experience with interactive archaeometry tools in digital simulations

Eduardo Zilles Borba<sup>1,2</sup>  · Ana Grasielle Corrêa<sup>1</sup> · Roseli de Deus Lopes<sup>2</sup> · Marcelo Zuffo<sup>2</sup>

Received: 18 June 2018 / Revised: 19 April 2019 / Accepted: 21 June 2019 /  
Published online: 12 July 2019

© Springer Science+Business Media, LLC, part of Springer Nature 2019

## Abstract

This is an experimental work presenting a study about usability experience of users in a cyber-archeology fully immersive 3D environment in Virtual Reality (VR). Particularly we research how users explore a realistic 3D environment with VR devices through archaeometry conventional techniques for archeology analysis and discoveries. Our main objective is to evaluate user's experience with interactive archaeometry tools of a population called archeologists, not a VR expert, but expert on the context of the real experience of practicing archeology in remote site; and compare the results with another population called VR experts, in this case, not experts in archeology, but in VR technology and multimedia applications. Several standard metrics will be used to collect data about their interactions with the cyber system (efficacy, efficiency, satisfaction, level of presence and cyber-sickness). Two hypotheses will be tested with this experiment: a) it is possible to represent the virtual world as realistically as the real one, in such way that a person unfamiliar with this kind of technology, in this case the archaeologist, can develop analytical process of discoveries in the VR model; and b) if this VR model is passive of exploration, virtually it is possible to create analytical tools that will help the archaeologist to manipulate archaeometry tools. Both sample population had participated in usability tests and the results are promising.

**Keywords** Usability · Virtual reality · User experience · Human computer interaction and multimedia · Telepresence

## 1 Introduction

Big data is a widely used term for naming complex or very large data sets. Challenges in this area include: analyzing, capturing, searching, sharing, storing, transferring and viewing data

---

✉ Eduardo Zilles Borba  
ezb@lsi.usp.br

information. Technological evolution in this field of research enabled the development of instruments and techniques for the acquisition of 3D data in different scales, depth and precisions. For example, 3D digital scanning techniques, such as LIDAR and stereo photometry, have reached such a degree of maturity that allows the scanning of archaeological sites in 3D at a very high resolution [31].

Archeological sites, especially the ones with rock art, are very fragile and susceptible to destruction [10]. It means, when archaeologists start the digging procedure, the original information of the landscape may never be recovered. So, archaeological sites are ephemeral spots. After all, the site is exposed to the environment and, consequently, undergoes the process of natural degradation, besides being susceptible to destruction by vandals [12]. Having the digital record of these sites could be a way to preserve them for the future [6].

For those reasons we recreated, with high fidelity on images simulation, the archeological site of Itapeva Rocky Shelter, in São Paulo (Brazil), through modern techniques of 3D data capturing. The visualization of the virtual environment (VE), which we called cyber-archeology, can be made through a high definition head-mounted display set (*i.e.* HTC Vive or Oculus Rift), allowing the user to feel a great sensation of immersion in the VE [4]. In addition to the visual perception realism, we have created three virtual archaeometry tools for user interactions with the 3D environment: the annotation tool, the rock painting tool and the artifact classification tool. All those 3DUI tools allow the user to perform scientific research activities using a head-mounted display (HMD) and a 3D input device (SteamVR joysticks) [35]. Our idea is to enable archaeologists to make an exploratory analysis of the acquired data and even measure the degradation of the field. Also, this application urges as an education tool for graduation students in the archaeology field, which would not be able to explore such an ephemeral location, but in VR could feel like what it is like to be in a field activity [35]. From the museology perspective, cyber-archeology would also allow the lay visitor to have an initial idea of spatiality that cannot be transmitted through photography or videos (forms, scales, proportions, textures, lights and shadows) [6]. From the scientific point of view, it is possible to discuss and analyze technical aspects, such as: to observe the state of the paintings on the wall, to discuss about where excavations were made, to reconstitute paintings or, even, to get extra data about the artifacts (size, weight, classification) and others [12, 31].

Despite the advances in the cyber-archeology field, the potential impacts on user interactions with immersive archaeological VR models are still unclear. So, studies about it must be carried through. In our perspective, it is required attention for the application of usability evaluations, especially by involving users with non-conventional interaction technology devices. To conduct such a research about design and interactions with VR environments, the international standards and usability test methods are available to guide us in this article [7, 15, 16, 22–24, 30]. It means, based on our ArcheoVR application, this paper proposes the conduction of an exploratory study with users, who report their experiences with the cyber-system (the interface usability, the appropriation of interaction devices, the sense of presence and immersion, cybersickness, etc.), helping us to evaluate its performance validating and/or suggesting reconfigurations on the application conception itself.

In short, the main objective of this paper is to evaluate user's experience with the interactive archaeometry tools of a population called archaeologist (not a VR expert, but expert on the context of the immersive experience) and compare the results with another population called VR experts (in this case, not experts in archeology). Two hypotheses will be tested: a) it is possible to represent the virtual world as realistically as the real one, in such way that a person unfamiliar with this kind of technology, in this case the archaeologist, can develop her/his

analytical process of archaeological discovery in the VR model; and b) if this VR model is passive of exploration, virtually it is possible to create analytical tools that will help the archaeologist to manipulate archaeometry tools. Ten people participated in the usability test and the results are promising.

## 2 Previous work on electronic archeology simulation

In archaeology, excavation is the exposure, processing and recording of archaeological remains [12]. It may be understood as a work of different scales: intra-site (a single space) or inter-site investigation (multiple sites and landscapes). Both kind of investigation could take a long time to be accomplished (from month to decades). Despite adopting rigorous procedures, the excavation process is a destructive one and once a place is excavated the original physical site is no longer available [11].

On this perspective, during the IEEE VR Conference 2016 we presented a 3D immersive and interactive environment named ArcheoVR [3, 5]. Its proposal was to allow the general public to visually and explore an archaeological site located in São Paulo (Brazil) called Itapeva Rocky Shelter (Fig. 1). In this opportunity, the VR model was coupled with visual analytics and editing tools which optimize the exploration, harvested of modern archaeometry digital techniques (combining LIDAR and stereo photometry point-clouds) [32]. During the journey the user was allowed to navigate and interact with pre-defined digital objects using control devices (keyboards, mouse and Razer Hydra joystick). To visualize the 3D scene in a realistic way user had to wear HMD device (Oculus Rift DK2).

It is important to underline concepts that guided us to design this version of Itapeva 3D VR model, such as: cyber-archeology, presence in VR and gamification narratives. All technical process of developing this application with LIDAR and 3D image mapping and modelling a huge point-cloud of 3D mesh can be confirmed in our full papers published in IEEE and Siggraph conferences [4, 35].



**Fig. 1** User exploring the 3D Itapeva VR model with HMD device

Doneus et al. [10] claim that cyber-archeology field is a natural evolution of archeology itself. Forte [12] agrees, claiming that cyber-archeology combines the state-of-the-art of virtual archeology and e-Science. Indeed, cyber-archeology has generated discussion in several areas. Since the precision of details transposed from the physical to virtual world until the possibilities of reconstructing and re-exploring the past through VR and its impacts on the archeological knowledge acquisition process.

Forte [12] suggests that in a VE it is more appropriate to talk about the simulation of the past rather than the reconstruction of the past. Same author talks about the importance of embodiment and the power of VR models as tools to augment the process of validation on the entire cognitive process of the archaeologist.

Previous work has also discussed about how VR support and produce higher feelings of presence in historical exploration applications. For example, [6] say that, when the main objective of a simulation is to save and/or to preserve past experiences, immersive media indubitably stands out as much more effectiveness than visual and audiovisual media (photography, video, illustration and others). Indeed, multi-sensory interfaces should allow the user to feel how it was the aura, the noise, the crowd or the atmosphere of historic scenes (*i.e.* explore the Jurassic Era, walk through the Great Wall of China, see through a first-person perspective scene the World War I, and others).

### 3 Virtual continuum in cyber-archeology

Virtual Reality (VR) is an advanced interface model that allows user to view, to interact and to manipulate digital contents through a computer in a natural way. In general, supported by electronic devices (3D display, sensors, cameras and others) it creates an effect of physical reality in the synthetic scenario, stimulating the user immersion in the artificial context and, therefore, creating the idea of presence in another reality which is not the physical one [26]. From the perspective of [8] or [36] three pillars support the VR experiences: realism, interactivity and engage (Fig. 2).

Realism indicates the capacity of the virtual environment to simulates plastic, aesthetics and surrounding elements as their original versions (forms, colors, scales, proportions, noises, voices, etc.). Interactivity reveals how the user interact with the digital interface. If more similar to the real-world operation and interactions with spaces and objects, more immersive would be her/his notion of interactivity (to walk, to talk, to touch, to catch, to manipulate, etc.) [13, 33]. On its way, the engagement is totally associated to the production of a wide perception of space and, also, to the attention transfer phenomenon. It means, the engagement

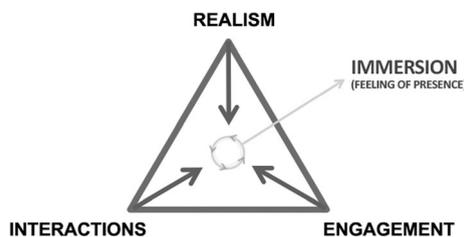


Fig. 2 The three pillars of VR immersive experiences

considers the 360° space, which produces the notion of a scene around the user (much higher than one generated by a computer flat screen). But, also, it may be considered the narrative elements directly involved with user attention transfer, because the plot capacity of catching user attention related to her/his preferences, memories, etc. (storytelling, gamification and others) [27].

Anyway, VR has a huge variety of environment systems. To classify those technological environments through a functional taxonomy, [21] developed a continuum in which different environments are separated by its proportion of real and virtual components. This classification was called the Virtual Continuum (Fig. 3).

We may understand this continuum as a scale between the completely virtual scenario (the binary space) and the real one (the physical space). Between virtual and real there are mixed realities where environments could be evaluated by more or less presence of real or virtual components. For example, the Augmented Reality (AR) has virtual objects complementing the real world (*i.e.* Pokémon Go, Microsoft HoloLens; QR code media revealed) [1]; and the Augmented Virtuality (AV) has real spaces and/or objects complementing the user experience with virtual scenarios (*i.e.* Nintendo Wii, VR Thor) [9].

The virtual continuum theory is important for this study because it allows us to think about the possibility of creating a continuum of the archaeology research through VR models. It means, through the virtual data collected with a complex methodology it may be possible to generate a new paradigm to the archaeologist work (Fig. 4). After all, the archaeologist would be able to explore the synthetic model in a similar experience of the real-world. In this case, VR exploration emerges like a continuous flow of the archaeologist work: the archaeologist in the loop. This virtual loop in cyber-archeology was presented in our past work [35], where we explained how to collect data from the real-world using equipment as LIDAR, 360° imaging and planar imaging to generate a huge point-cloud register (a 3D mesh) [32]. This digitalization process, according to [4, 35], is fundamental to create a photorealistic VR scene through

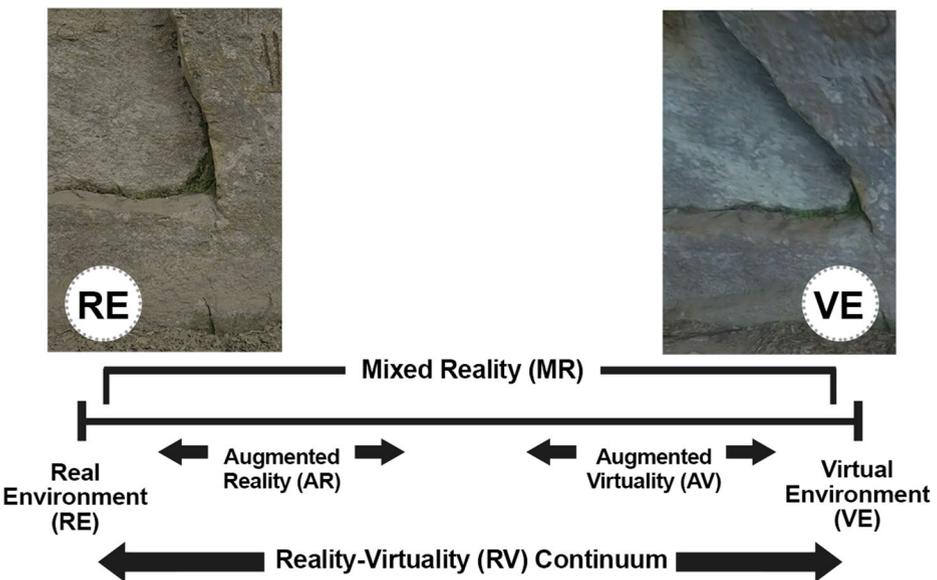


Fig. 3 Virtual continuum scheme applied in the archeology real and virtual environments

computer graphic resources. Looking for the block diagram of the interactive excavation process of an archeological site in VR we may underline this study is totally focused on the user experience (the archeologist) with the ArcheoVR cyber-system we have created in the past year. For example, we want to understand the efficiency, efficacy, satisfaction, levels of presence of users when interacting with the digital archaeometry tools, such as: annotation, classification, image analysis and marking hotspots.

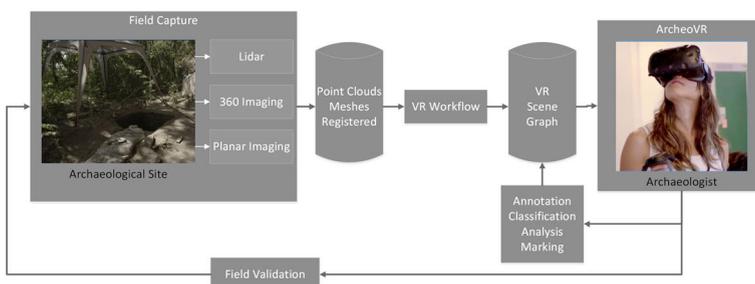
It is not just about a virtual tool, but an extension of the real-world activities [20]. The digital model allows the user to keep analyzing rock paintings in the lab facilities, not through photo or video, but as if she/he was really transposed to the natural field. To reflect about this real-virtual transposition we suggest a pseudo-code representing the interactive excavation of the archeological site in the real environment (Fig. 5).

#### 4 Overview of the usability testing in VR simulation scenarios

Usability is defined by ISO 9241-11 as “the extent to which a product can be used by specific goals with effectiveness, efficiency and satisfaction in a given context of use” [16]. Usability tests are commonly used in usability engineering and human computer interactions (HCI) disciplines. The main propose is to verify how easy a product can be understood and manipulated by the user.

Usability of a product can be evaluated through different methods. The heuristic evaluation, introduced by Nielsen [22, 23] is the most common method. It does involve experts evaluating the interface based on a set of usability criteria or heuristics. In other way, usability tests involve real users evaluating the product and/or the experience with a digital interface. Its objective is to describe the performance of a typical user in a carefully prepared experience with predefined tasks for which the product was intended.

As Nielsen said [24], the real advantage of usability test is that technical resources needed to conduct it are simple and modest. It means, even if a high number of volunteers may present a great study, a small sample of five users can identify 85% of the usability issues [24]. Also, this kind of tests can be performed either in a controlled environment (*i.e.* lab) or in the field – but in this case interruptions and noises may occur generating influences on the final results. Both cases allow researchers to gain an understanding of how a product can be used in a specific context. To perform the usability test and evaluate the product performance, [30] recommend a series of activities planned to obtain effective results: a) to design the experience, b) to collect data, c) to interpret data and d) to report the results. During the



**Fig. 4** Block diagram of the interactive excavation process of an archeological site in VR

```

Archaeological_excavation_protocol {
  Set Scientific_Question;
  Set Funding;
  Plan_excavation(Scientific_Question); /* location, squares, depth, methodology, ... */

  while (Scientific_Question is pending or Funding available) {
    Excavate(Depth_Step);
    Collect_and_Check_soil_sediments (Soil_samples);
    Collect_artifacts(Artifact_list);

    For each Artifact collected {
      If (Artifact is relevant) {
        Classify(Artifact); /* Lytic, Fauna, Flora, Ceramics, ... */
        Register(Artifact); /* annotate, coordinates, depth, picture,    archeologist, ... */
        Store (Artifact); /* for further analysis, dating ... */
      }
    }
  }
  Close_excavation();
}

```



**Fig. 5** Pseudo-code representing the excavation in the real-world

design preparation it is imperative to define tasks the users will execute. Tasks, in addition to being typical, should be related to the objectives and main functions of the system. It is advisable to create short, clear and objective scenarios in order to facilitate the tasks accomplishment by the users and, thus, to increase the effectiveness of results, since the scenarios are stories representing real situations of use and allow to create a more realistic environment, eliminating the superficiality of any test. Still, in the preparation phase it is necessary to select and recruit people, having as criterion the profile and characteristics of each user. The location planned for testing should be prepared to receive participants with comfort, privacy and safety. The equipment must be tested before the tests. To do so, it is necessary to perform a pilot test to configure and verify the operation of the digital environment, in addition to checking and validating the test plan [23].

The next step in usability tests is data collection. It involves pre-test, questionnaires, observation sessions and/or post-test interview. The observation should be conducted with annotations and camcorders that capture both: the actions of the users (even their facial expressions) and testimonials. The think aloud method is recommended as well:

“listening to a user’s thoughts enables them to understand the reason for their actions and this information is valuable in the testing process” [24].

The pre-test questionnaires should be applied before starting the empirical experience with the sample population. They are imperative for collecting information about user’s profile and demographic data. The post-test questionnaires should be applied after the interaction sessions and aim to diagnose satisfaction, impressions and subjective perceptions about the experience.

In the data interpretation and consolidation of the results, the obtained data are analyzed to obtain quantitative and qualitative results. A quantitative focus allows the researcher to test hypothesis, discover trends, compare alternative solutions and verify if the system usability meets the required goals recommended by ISO 9241-11. The qualitative focus requires a joint and interpretative analysis of all data collected, allowing the identification of the origin of any problem, explanation of problems or justification and validation of applied solutions.

Finally, the last activity of the usability tests is to present a report of results containing: objectives and evaluation body; brief description of usability testing methods; number and profile of evaluators and participants; tasks performed by the users; graphs and tables illustrating the measurements obtained; list of problems; and, also, the suggestions for solutions [24].

## 5 Materials and methods

Following it is presented the methodology, tasks, materials and settings applied to conduct the empirical study with volunteers who interacted with the 3D Itapeva Rocky Shelter VR Model.

### 5.1 Methodology

This work conducts an experimental study presenting scientific results of the user experiences with a cyber-system in VR environment. It is important to underline that it presents more like a technical report based on archival attitude than a technical-centric discover in the cyber-archeological field. To accomplish the objective of evaluating user experience with interactive archaeometry tools in a digital simulation, we suggested the collection of a series of data related to the user experience and usability perception when exploring VR environments. It means, we adopted an empirical and exploratory methodology with users exploring and interacting with the technical interfaces/3D simulation systems in our lab [23, 30]. Metrics composing this evaluation about usability in VR involved efficacy, efficiency and satisfaction [7, 22, 24], level of presence and immersion [2, 14, 19, 25, 28, 29], cyber-sickness [17, 18] and others.

### 5.2 Study sample and settings

To participate in the experiment the volunteers were separated in two groups: the VR experts, not specialists in archeology (Group A); and the archeologists, not experts in VR (Group B). To create the Group A users were randomly selected from the invitation on social networks inside the university campus. The participants of Group B were selected by invitation from an expert in archeology (Prof. Astolfo Araujo, lead Archeologist in the Museum of Archeology and Ethnography at University of Sao Paulo). The inclusion criteria to compose our sample were:

- **Group A:** (1) VR experts; (2) without any knowledge about archeology
- **Group B:** (1) professional archeologists or students, without any knowledge about VR.

A pilot test was conducted with an archaeologist to validate our test protocol and also to estimate the total time spent in training (learning phase) and in the execution of required tasks. After this free session and informed consent, participants were scheduled to participate in the usability study conducted in the usability lab equipped with tables, chairs, notebook computer, HMD (HTC Vive), 3D input device (SteamVR) and a video camera.

### 5.3 Instruments and outcomes measures

Basically, we used demographic and IT/VR knowledge and usability measures.

#### 5.3.1 Demographics and IT/RV knowledge and experience

A pre-tested questionnaire was applied at the beginning of each session to assess the volunteers experience with VR or digital game narratives [6]. They were asked about their age, gender, education degree and previous experience with videogames or VR devices, including here the frequency of use (Group B) and/or the frequency of visits to the archaeological site of Itapeva (Group A).

#### 5.3.2 Tasks and scenarios

Tasks of the training were defined using the system's functionalities. They were structured in the following order: creating points of interest; creating text notes; creating painting; editing text notes; editing paintings; using the teleportation tool; removing text notes; removing the paintings; deleting point of interest; accessing the transition visualization (a tool to compare the computer graphics scenario with a 360° photography of the same one).

The usability assessment tasks were listed from scenarios. Texts and paintings were defined *a priori*, so users would not be distracted by thinking about what they would type or paint (this could interfere with task time). Six scenarios were created step-by-step:

- **Scenario 1 (creating artifacts):** create a point of interest named *point 1*. Inside it creates a text note named *register 1* and enter the text *register rock*. Then create a painting of a man's toothpick in blue color and name it *paint 1* (Fig. 6).
- **Scenario 2 (editing artifacts):** change the contents of *register one* to *rupestrian register 2015*. Then redo the *paint 1* drawing the same toothpick but in red with the stroke thickness in 25 mm and depth 0.8 mm (Fig. 7).
- **Scenario 3 (moving around the scene):** move yourself to the location behind the big rock. Then, return to your point of origin.



Fig. 6 Creating a point of interest on the wall

- **Scenario 4 (creating new artifacts and point of interest):** use the teleport command to create another point of interest behind the big rock and call it *point 2*. Inside it create a text note called *register 2* and enter the text *rupestrian register 2016*. Then, create a painting of a man toothpick in yellow color and name it as *paint 2*.
- **Scenario 5 (deleting artifacts):** observe that you have created two points of interest on the wall. So now, delete the *register 2* inside the *point 2*. Finally, delete *point 1* and *point 2*.
- **Scenario 6 (transition):** make the transition from the virtual (computer graphics) to the real scene (360° photography). Take a look all over the place. When you're done, just go back to the virtual scene and move to your point of origin.



Fig. 7 Editing stroke thickness of the painting tool

## 5.4 Usability measures

Established by ISO 9241-11, the usability metrics (efficacy, efficiency and satisfaction) were studied from the data collected with the questionnaires.

Efficacy were calculated through success in completing tasks and by number of errors occurred during interactions. For this specific work, the effectiveness was measured by the conclusion of the task, being assigned the following values: (1) *completed with ease*, when the user was able to perform the task without any help from the moderator; (2) *completed with difficult*, when the user performed the task but asked for help or tips from the moderator; (3) *not completed*, when the user was not able to complete the task, even with minor moderator tips. An error was coded when the subject performed any kind of errors (*i.e.* not solving a task).

Efficiency included the level of effort and use of resources by the user to achieve usability goals. This is typically measured by task execution time. So, efficiency is related to the execution of tasks by users in the shortest time possible, thus,  $t_{med}$  must approach  $t_{min}$  (less time spent in performing tasks) to indicate a higher level of efficiency of the product. In this work, the efficiency was determined by measuring the time (in seconds) of each task performed individually and calculating the average time of each task considering all users. Task time began to be timed when moderator verbally finalized the task request and ended when the user verbally said the word: *ready*. The user was always encouraged to think aloud while performing tasks.

Task performance measures (effectiveness and efficiency) can be used as indicators of presence in virtual reality environments [13]. The characteristics of the user, such as capacity and motivation, also influence the performance of tasks. For example, a lot of studies associate user presence feelings to the effectiveness and efficiency of task performances [2, 14, 19, 25, 28, 29, 34].

In this study, we used the Slater-Usoh-Steed Questionnaire (SUS Questionnaire) [25]. It comprises six questions, where each answer must range from 1 to 7, from the minimal extreme (1 = *low presence*) to the maximum extreme (7 = *high presence*). Each one of the six issues involve aspects of presence, such as the feeling of being in the virtual scenario, the believing of virtual scenario becoming a physical reality or the remembering of the experience as a place visited more than images visualized with a HMD device. In this way, each of the six questions involves a scale of 1 to 7, where the highest scores report to higher levels of presence.

The satisfaction of the interaction was measured through the System Usability Scale (SUS) questionnaire, one of the best known and simplest methods of ascertaining the usability level of a system. The popularity of the method is due, among other reasons, to the fact that it presents an interesting balance between being scientifically accurate and, at the same time, not being extremely long for the user or the researcher. Criteria that SUS helps evaluate (from the user's point of view): effectiveness (can users complete their goals?); Efficiency (how much effort and resources are needed for this?); Satisfaction (was the experience satisfactory?). The questionnaire consists of 10 questions, each of which responds to the 5-degree Likert scale, where 1 means *Strongly Disagree* and 5 means *Strongly Agree*. The score is calculated as follows: subtract 1 from the odd answers; Subtract 5 from even responses (2 and 4); Sum all the values of the 10 questions and multiply the result by 2.5. This is the final score that can range from 0 to 100. The mean of SUS is 68 points, a value lower than this means that the product has serious usability problems. We also used the Simulator Sickness Questionnaire [17] to measure symptoms of cybersickness [18]. In cybersickness syndrome, users are given visual cues that can induce an illusion of body movement while their bodies are physically

stopped. This would cause disagreement between visual, vestibular and proprioceptive stimuli, culminating in symptoms. These disagreements may be due to a range of factors such as: error in tracking the user's position in relation to the program, delay in updating the body position, tremor or oscillation of the represented body parts, distorted graphics, poor optics, image flickering. Illusion of the movement itself, the delay between the time that a physical movement is performed and the response time of the computer with modification in the generated image, the field of vision, individual variables and amount of head movements are also pointed out as Influence factors for the disagreement between stimuli. Factors related to oculomotor and nausea were classified as absent, mild, moderate or severe and scored from 0 to 3, respectively. The study of cybersickness is important for implementing system improvements, to help preserve user well-being and reduce abandonment to virtual exposures. All video recordings were carefully analyzed in order to encode the specific metrics defined above: success rate, error and time of each task. The session ended with a discussion of the users' impressions about the experience they had just completed and filling out the questionnaire: a) SUS of Presence (6 questions); Interaction SUS (10 questions) and Cybersickness (8 questions). Each test had average duration of 30 min per participant (10 min for the presentation and training; and 20 min for the execution of the tasks). Each participant received a toast (chocolate) after the end of the session.

## 5.5 Data analysis

The collected data was tabulated in the Microsoft Excel worksheet. Descriptive statistics, averages and standard deviations were calculated to aid in the analysis of efficacy and efficiency of tasks, satisfaction of the interaction, sensation of presence and symptoms of cybersickness. Results among groups were compared with sociodemographic data to distinguish differences and similarities between user profiles and their performance.

## 6 Results and discussion

Following we present the results and discussions about the data collected during the empirical experience with ten volunteers in the cyber-archaeology VR model.

### 6.1 Population

Between 10 participants, 5 composed the Group A (VR experts) and 5 composed the Group B (professional archeologists). They present the following profile:

- **Group A:** all males, with average 33 years-old and full graduate studies; rarely use electronic games, but when using they prefer computer platforms than game consoles; all consider themselves experience or very experienced with the use of VR technologies because with high frequency use interactive VR environments in their research labs; only one of them have visited the archaeological site of Itapeva.
- **Group B:** three were females and two males; all with an average 34 years-old, with complete graduate studies; rarely or never use computers to play, but eventually use smartphones for this purpose; have little or no experience with game consoles; all of them

already have used 3D stereoscopic goggles in cinemas, but they had no experience with interactive immersive VR environments; all volunteers have visited the archaeological site of Itapeva at least once in 2016.

### 6.2 Effectiveness

The Fig. 8 shows the results of effectiveness of tasks performed by Group A. Teleporting (scenario 3) and transiting between environments (scenario 6) were the easiest to accomplish, being easily completed by all users. The tasks of creating artifacts (scenario 4) and deleting them (scenario 5) were the most difficult to complete. Respectively, 26% and 32% of the users failed to complete these tasks, presenting major difficulties in handling the SteamVR to create or delete an artifact in the scene. This may happen because, in scenario 4, some users did not teleport before creating the new point of interest. So, at the time of creating the note or the drawing they could not see the information clearly. After all, their avatar body wasn't close enough of the point of interest created on the wall.

Another problem identified during the tests was that users confused the rename function (*i.e.* renaming an edit note) with the edit text notes function (*i.e.* creating a text inside the note). It means, the difference between the function of changing the file's name and the function of inserting content inside it did not seems so obviously to the users. It is a point to be improved in the VR environment, on the way to make it clear to any user.

The error rates reflect the difficulties of accomplishing tasks. Fails included: a) not using the teleportation to create artifacts in distant zones (point of interest); b) confusion between renaming the notes field and editing text notes inside it; and c) deleting point of interest instead deleting only a text note or a painting. There was only one failure that prevented one of the tasks from the scenario 5 from continuing: one user excluded the point of interest rather than deleting the text note.

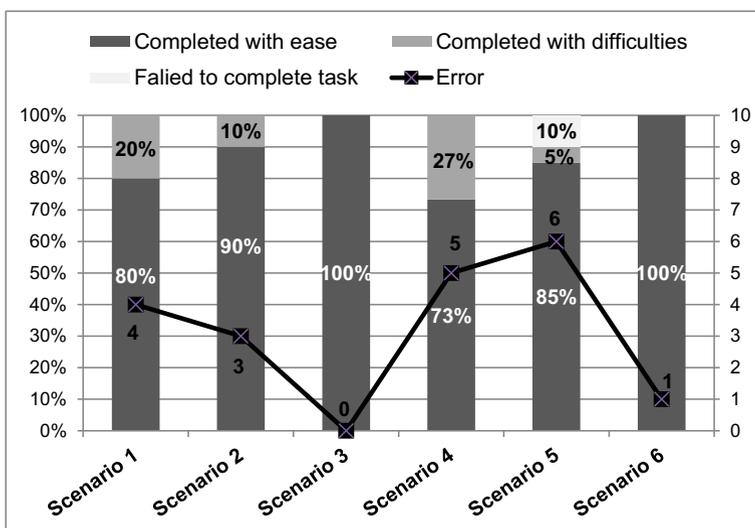


Fig. 8 Effectiveness of the tasks performed by Group A

The Fig. 9 shows the results of effectiveness of tasks performed by Group B (professional archeologists). Similar as Group A tasks of moving around the scene using teleportation (scenario 3) were the easiest to accomplish, being completed easily and with no error by all users. Creating artifacts (scenario 1), editing artifacts (scenario 2) and removing artifacts (scenario 5) were the most difficult to complete with, respectively, 35%, 25% and 22% of users fails. Difficulties accomplishing the tasks were: a) not remembering how to create artifacts; b) confusion using the virtual keyboard, since SteamVR automatically reverses the letter selection functions by the control (from the right to the left hand); c) confusion using the function rename and edit text note; d) do not remembering to select the pencil icon before scratching/painting the wall. Similar to the Group A, two users deleted the point of interest rather than deleting one of their artifacts, which unfortunately prevented the completion of the task.

By comparing the groups, we can say the tools of archaeometry (annotation and painting) were usable in the cyber-archeology field, except for some adjustments in the interface.

Two major problems were identified in both groups: a) the confusion between functions of renaming the note file and editing texts inside it; and b) deleting a point of interest on the wall (unintentionally) due to the large distance of user position to the point created on the wall. Therefore, it will be necessary to redesign these two functionalities of the system. The rates of tasks performed were high, which proves the first hypothesis: it would be possible to represent the VR world as realistically as the real one, in such way that a person unfamiliar with this kind of technology, in this case the archeologist, can develop her/his analytical process of discovery in the VR model.

### 6.3 Efficiency

Looking to the Table 1, it is possible to identify that creating artifacts (scenario 1), editing artifacts (scenario 2) and creating new artifacts (scenario 4), performed by Group A, consumed as much time as expected. It may be explained because these tasks involve typing (insert texts)

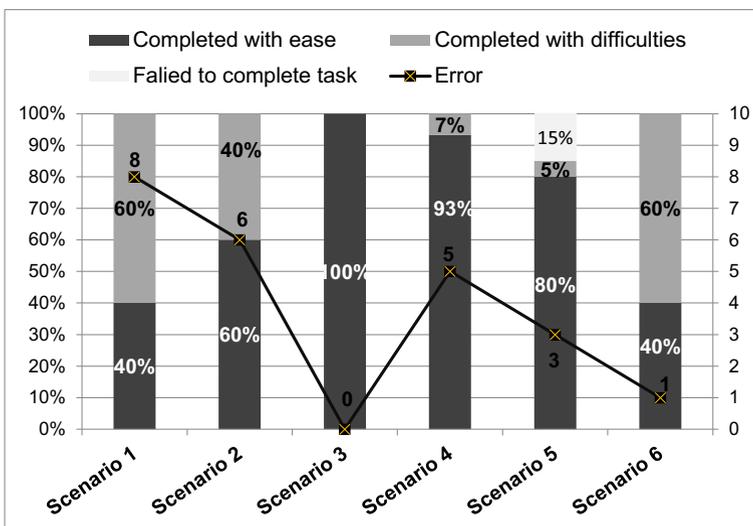


Fig. 9 Effectiveness of the tasks performed by Group B

**Table 1** Scenario success and error rate by task

SCENARIO			1	2	3	4	5	6
A	Time per Scenario (seconds)	Mean (SD)	33,6 (13,9)	48,3 (25,2)	13,8 (3,5)	36,9 (16,5)	10,2 (4,5)	18 (3,6)
	Range		11–59	16–102	10–20	14–70	3–20	15–25
B	Time per Scenario (seconds)	Mean (SD)	61,1 (31,3)	49,5 (18,6)	13,8 (1,9)	47,2 (21,3)	11,3 (4,1)	42,6 (18,6)
	Range		12–119	24–82	12–17	24–93	5–18	1–72

and drawing actions that demand a greater effort. On the other hand, tasks of moving around the scene (scenario 3), deleting artifacts (scenario 5) and transition from the virtual to the real 360° photo (scenario 6) were the quickest.

In the Group B we have identified a similar behavior. It means, tasks in scenario 1, 2 and 4 were the most time-consuming due the actions of editing notes and drawings, aggravated by the difficulty of this group to remember the difference between renaming text field and editing text notes (there was a significantly interference from the moderator to help users to accomplish those tasks). The delay in completing tasks in scenario 6 was due to the difficulty of remembering how to use the transition functions and also because the archeologists spent more time observing and comparing the landscape details (the computer graphics versus the 360° photo).

#### 6.4 System usability scale (SUS)

The average score of SUS interactions for the whole group (A + B) was 78.5 (Fig. 10). It indicates a good satisfaction among users (archeologists and VR experts) – as seen in Fig. 10. However, there were variations in the scores with a minimum value of 70.0 and maximum of 97.5 (a range variation of 27.5 points).

#### 6.5 Slater-Usoh-Steed questionnaire (SUS Questionnaire)

When observing Fig. 11 it is notable that both groups attributed high scores to the questions related to the perception of presence in the VR context. After all, on a Likert-Scale of 1 to 7, no average response for the six questions was below the 4.6 score.

As expected in Group A there was a higher requirement of users to dive into the experience. That is, because they are experts, analyzed technical details related to the immersion in the scene (the sensory experience induced by the technological equipment). *i.e.* they looked closely at the aesthetic details of landscape composition and modes of interactions with objects in that landscape. So, when looking to the question 1 (*How much did you feel present in the scenario presented?*), question 4 (*During the time of your experience, was it stronger to be present on the stage or to be somewhere else?*) and question 6 (*For some time during the experiment did he think he was really on the scene?*), it is great for the system to identify answer with 5.4, 4.6 and 5.2 points. The average score 5.0 for the question 2 (*To what extent has the scenario presented become a reality for you to forget the laboratorial context in which the experience takes place?*) reveals the users had an elevated presence feeling in the VR model. Indeed, with they had headphones, probably the lab environment would be more neutral to the experience and then the, consequently, this score could reach 6 or 7 points in the

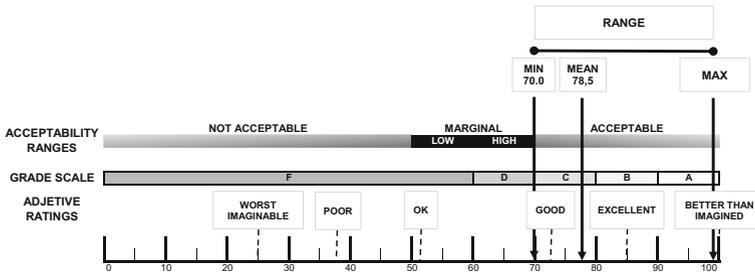


Fig. 10 Overview of modified SUS rating table with inserted value ranges

evaluation scale. Question 3 (*How much do you remember the scenario, do you see it only as an image or as a place where you were present?*), is a controversial issue and also difficult to explain to test participants. However, 3 of the 5 participants in the VR expert group pointed out that they remembered more of a place they visited than an image they saw, claiming that the notion of forms, scales, perspectives, depths, textures, leftovers and enlightenment transported them to a true space notion. Reproduced from a space. Perhaps the final average score on this answer was not only higher because one of the users felt that their experience was entirely with an image and not with a site they were on. This same user was also very demanding in previous issues about feeling present in the virtual context. The final average score for this question was 4.6 points. At least, question 5 (*To what extent have the structural conditions - colors and objects forms - of the scene reminded you of other similar situations in which you have been?*) presented a satisfactory 5.2 points, probably because the system was constructed in a very realistic way when comparing the forms, scales, depth and colors of the virtual objects with the real ones.

In turn, group B presented a more elevated sense of presence in the VR context created to simulate Itapeva Rocky Shelter. For all the questions the archaeologists scored higher indices to point their presence feeling in the virtual environment. For example, in question 1, 4 and 6 they presented, respectively, 6.0, 6.4 and 6.0 evaluations. It means, they really felt like being

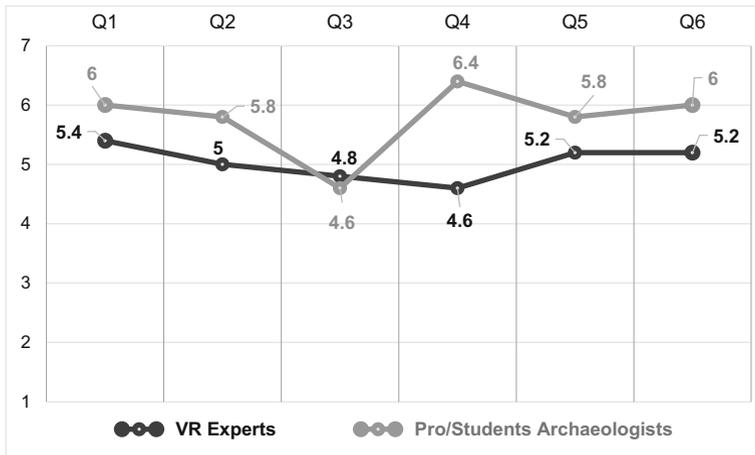


Fig. 11 Average evaluation on perception of presence in VR

inside the simulation. One of them said: *the feeling is really as being in the original place. The experience of the sphere (the view all over the place and the changing of computer graphics aesthetics to 360° photography) gives me the view from above of the site, so I felt a little vertigo, but it was very interesting. Congratulations.* Another commented: *the way we created points of interest on the rock panel (the wall) was simple and easy to control allowing a very similar interaction experience as in the real world.* About question 2 and 3, we underline a similar perception feeling identified when analyzing group with, respectively, 5.8 and 4.8 points scored. Anyway, in both questions, again, the archaeologists showed a little more perception of presence in the virtual context than the VR experts. At least, in question 5, the users of this group considered a very familiar visual experience with colors, forms and behaviors of virtual objects when comparing to the real ones. They scored a significantly 5.8 points in a maximum of 7.0 points. We believe this question is really imperative for this group, after all they are the people who have being in lots of time in Itapeva site and have an extremely sharp look at the details of visual and plastic composition of the original landscape.

As final thoughts about the presence feeling we may say that mostly of the users of both groups had a high feeling of being in there (Itapeva’s Rocky Shelter) than exploring a set of images. Also, believe the combination of visual realism, user-friendly interactions and the sense of being involved by the scenario (a 360° perception of the scene) stimulated the users to believe they were in the original archeological site, at least during a few moments for the VR experts and with more intensity for the archaeologists.

### 6.6 Sickness questionnaire

As observed in Fig. 12, the register of nausea and oculo-motor related symptoms were relatively low. Group B (archaeologists) presented symptoms of vertigo at a moderated level. While, Group A (VR experts) presented slight or none symptoms related to oculo-motor, among which the tired sight and the blurred vision were the most frequently mentioned. Curiously, the Group A was the only one indicating a severe nausea problem in the user experience.

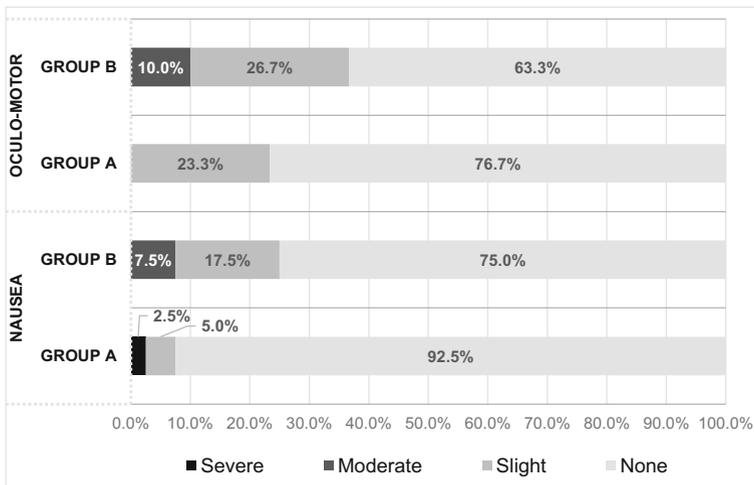


Fig. 12 Average rating for simulation sickness

About nausea we may say users had not notice a huge problem. Indeed, the Group B, composed by archeologists, registered 7.5% of moderated, 17.5% of slight and 75% none nausea feeling. As expected, the Group A, composed by VR experts, almost did not complain about nausea. The numbers show an impressive 92.5% of users who did not have any trouble exploring the virtual world with the HMD device. After all, they are used to wear VR devices for long hours for developing and testing 3D simulation projects. But, interesting, among the Group A we registered 2.5% of severe nausea symptoms. It represented a very small number of the experiment population, but it still got our attention. The user complained about sweating and dizziness with eyes opened while moving around the scene (scenario 3). In general, we must underline that nausea complain were related to vertigo and sweating (groups A and B), general discomfort feeling (group B) and/or dizziness with eyes opened (group B and only one user of Group A).

By it way, the oculo-motor symptoms were highlighted by both groups of users. The group of archaeologists presented 10% indices for moderate problems, 26.7% for slight and 63.3% for none. The main observation about the sickness were related to eyestrain, difficulty maintaining focus, blurry vision and difficulty concentrating. While the VR expert group did not present any severe or moderate problem with oculo-motor cyber-sickness.

## 7 Conclusion and future work

This study demonstrated the application of systematic usability methods during cyber system interactions. Our results showed how the usability standard of ISO 9241-11 can be used for a set of efficiency, efficacy and satisfaction measures. We also analyzed the level of presence in the immersive virtual environment and issues related to cyber-sickness (nausea and oculomotor).

Also, we may have concluded that it is possible to represent the virtual world as realistically as the real one, in such way that a person unfamiliar with this kind of technology, in this case the archaeologist, can develop her/his analytical process of discovering in the VR model (Hypothesis 1 = true). Data collected with the ten users proved it through results achieved with the presence questionnaire (the immersion feeling of really being in the Itapeva site and not in anywhere else). It means, the perception of being there was achieved with the immersive HMD equipment, creating a realistic visual experience (realistic forms, scales, perspectives, textures, colors, lights and shadows). The results achieved with the VR expert's opinion - who are much more critics about the immersion and presence feeling in VE - prove that point.

As final conclusion we can say the data collected by the usability tests revealed us the potential of the analytical created by our interactive tool with the 3D input device (text note, paintings, teleportation, etc.). So, yes, the VR model is passive of exploration (Hypothesis 2, part 1 = true). Also, the usability test presented important data about simulation of behavior and actions of an archaeologist in the field. It means, it is clearly possible to create analytical tools that will help the archaeologist to manipulate archaeometry tools through a VR model. Anyway, we must to agree the interactive tool need some improvements, especially in order to become a more simple and intuitive operation resource for archaeologists (Hypothesis 2, part 2 = true and false).

As future work we have three topics in our research agenda: a) to keep collecting virtual data from the real world excavation and transposing all the process of Itapeva's archaeological investigation to the VR model; b) to apply the improvement suggestions indicated by the users

with the objective of optimize the interactive tool and, consequently, augment the analytical process in the archeology field through VR techniques; and c) to create a multi-user experience for remote collaboration through telepresence tools; leveraging the realism and quality of a cyber-archeology application such as Itapeva 3D to allow geographically dispersed experts to analyze the data together can provide unprecedented insight into the data that has not been thus far possible due to lack of access because of geographical constraints [33].

**Acknowledgements** This work has been funded by Sao Paulo Research Foundation (FAPESP) and National Council for Scientific and Technological Development (CNPq-Brazil). Also, authors would like to say thank you to the archeology team led by Professor Astolfo Araujo from the Museu de Arqueologia e Etnologia at University of Sao Paulo (MAE-USP) for supporting us with archeological technical issues.

**Open access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

## References

1. Azuma R, Bailiot Y, Behringer R, Feiner S, Julier S, MacIntyre B (2001) Recent advances in augmented reality. *Computers and Graphics IEEE*
2. Barfield W, Weghorst S (2003) The sense of presence within virtual environments: A conceptual framework. In: Salvendy G, Smith M (eds) *Human-computer interaction: Applications and case studies*. Elsevier, Amsterdam, pp 699–704
3. Borba EZ, Cabral M, Lopes R, Zuffo MK, Kopper R (2016) A fully immersive virtual model to explore archaeological sites. *IEEE Virtual Reality Conference, Greenville (South Carolina, USA). IEEE VR 2016 Proceedings, Greenville: IEEE, v. 1, p. 326–49*
4. Borba EZ, Cabral M, Lopes R, Zuffo MK, Kopper R (2016) VR Model to Explore Archaeological Sites in a Non-destructive Way. *ACM Siggraph, 2016, Anaheim (California, USA). Siggraph 2016 VR Village Proceedings, Anaheim: ACM Siggraph, p.1.*
5. E. Z. Borba, A. Montes, M. Almeida, M. Nagamura, R. Lopes, M.K. Zuffo, A. Araujo and R. Kopper. “ArcheoVR: Exploring Itapeva’s archeological site”. *IEEE Virtual Reality Conference, 2017, Los Angeles (California, USA). IEEE VR 2017 Proceedings, Los Angeles: IEEE, p.463–464.*
6. Borba EZ, Zuffo MK (2016) Replay! When Past Experiences are Reassembled Through Virtual Reality. *IAMCR Conference Proceedings, Leicester*
7. Brooke J (1996) SUS: a quick and dirty usability scale. In: Jordan P, Thomas B, Weerdmeester B, McClelland A (eds) *Usability Evaluation in Industry*. Taylor and Francis, London
8. Burdea G (2003) *Virtual Reality Technology*. Wiley & Sons, New York
9. Cabral M, Belloc O, Montes A, Borba EZ, Zuffo MK (2016) VR THOR – Virtual Reality Training with Hotstick on Operations Risks. *IEEE Virtual Reality Conference, 2016, Greenville (South Carolina, USA). IEEE VR 2016 Proceedings, Greenville: IEEE, v.1, p.327–349*
10. Doneus M, Neubauer W, Studnicka N (2005) Digital Recording of Stratigraph Excavations”, *CIPA 2003, XX International Symposium, p. 451–456*
11. Forte M (2008) Digital Matter and Intangible Heritage. *International Journal Digital Culture and Electronic Tourism 1(2/3)*
12. Forte M (2011) Cyberarcheology: notes on the simulation of the past. *Virtual Archeology Review 2(1):7–18*
13. Friedberg A (2006) *The Virtual Window*. The MIT Press, Cambridge
14. Heeter C (2001) Reflections on real presence by a virtual person. *Proceedings of 4th International Workshop on Presence, Philadelphia*
15. International Organization for Standardization (1998) Ergonomic requirements for office work with visual display terminals (VDTs) – Part 11 Guidance on usability. *International standard, p. 9241–11. Geneve, ISO*
16. International Organization for Standardization (2010) Ergonomics of human-system interaction – Part 210: Human-centred Design for Interactive Systems. *International Standard, 9241–210. Geneve, ISO*
17. Kennedy R, Lane N, Berbaum K, Lillenthal M (1993) Simulator Sickness Questionnaire: an enhanced method for quantifying simulator sickness. *Int J Aviat Psychol 3(3):203–220*

18. Kennedy R, Lanhan D, Drexler JM, Massey CJ, Lilienthal M (1997) Comparison of cybersickness incidences, symptom profiles, measurement techniques, and suggestions for further research. *Presence* 6(6):638–645
19. Khanna P, Yu I, Mortensen J, Slater M (2006) Presence in response to dynamic visual realism: a preliminary report of an experiment study. *ACM Symposium on Virtual Reality Software and Technology*. Cyprus: ACM VRST, p. 364–367
20. McLuhan M *Understanding Media: the extensions of man*. McGraw-Hill, New York
21. Milgram P, Kishino F (1994) A Taxonomy of Mixed Reality: visual displays. *IEICE Transactions on Information Systems*, Tokyo, pp 1321–1329
22. Nielsen J (1993) *Usability Engineering*. Academic Press, Boston, p 358
23. Nielsen J (1994) Heuristic evaluation. In: Mack RL, Nielsen J (eds) *Usability Inspection Methods*. John Wiley & Sons, New York, pp 25–62
24. Nielsen J, Landauer T (1996) A mathematical model of the finding of usability problems. *Proceedings of the INTERACT '93 and CHI '93 Conference on Human Factors in Computing Systems*. ACM, Amsterdam, p. 206–213
25. Slater M, Usoh M, Steed A (1994) Depth of presence in virtual environments. *Presence: Teleoperators and Virtual Environments*. p. 130–144
26. Slater M, Wilbur S (1997) *A Framework for Immersive Virtual Environments (FIVE): Speculations on the Role of Presence in Virtual Environments*. Presence: Teleoperators and Virtual Environments, MIT Press, p. 603–616
27. Thon J, Millen D, Dimicco J (2012) Removing Gamification from an Enterprise SNS. *IBM Watson Research*, Cambridge, pp 1067–1070
28. Usoh M, Arthur K, Whitton MC, Bastos R, Steed A, Slater M., Brooks FP (1999) Walking>Walking-in-Place>Flying in Virtual Environments. *ACM Conference on Computer Graphics and Interactive Techniques*. SIGGRAPH, p. 359–364
29. Usoh M, Catena E, Arman S, Slater M (2000) Using Presence Questionnaires in Reality. *Presence* 9(5):497–503
30. Virzi R (1992) Refining the test phase of usability evaluation: how many subjects is enough? *Hum Factors*: 457–468
31. Vrubel A, Bellon O, Silva L (2009) A 3D Reconstruction Pipeline for Digital Preservation. *IEEE Conference*
32. Wei S, Cho S, Cho K, Um K, Won C, Sin S (2014) Traversable ground surface segmentation and modeling for real-time mobile mapping. *International Journal of Distribute Sensors Networks* 10(4):1–8
33. Wei S, Liu L, Tian Y, Sun G, Fong S, Cho K (2017) A 3D localization method in indoors environments for virtual reality applications. *Human Centric Computing and Information Sciences*. Berlin: Springer Berlin Heidelberg, p. 7–39
34. Youngblut C, Huie O (2003) The relationship between presence and performance in virtual environments: Results of a VERTS study. *IEEE Virtual Reality Conference*, Los Angeles (California, USA). *IEEE VR 2003 Proceedings*, Los Angeles: IEEE, v.1, p. 277–278
35. Zilles Borba E, Cabral M, Lopes R, Zuffo MK, Araujo AGM, Kopper R (2016) Virtual Excavations: using immersive technologies to visualize and interact with Itapeva's archaeological site. *IEEE Virtual reality Conference*, 2016, Greenville (South Carolina, USA). *IEEE VR 2016 Proceedings*, Greenville: IEEE, v.1, p.47
36. Zilles Borba E, Zuffo MK (2018) Déjà vu: revivendo experiências em ambientes de realidade virtual. *Interin (UTP)* 23(1):221–237

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Eduardo Zilles Borba** is Ph.D. in Communication and Information Sciences at University Fernando Pessoa, Portugal, in 2013. He is currently a Post-doctoral candidate in Department of Electrical Engineering and Digital Systems of the Polytechnic School at the University of Sao Paulo (USP), Brazil. Eduardo is also researcher in the Interdisciplinary Center of Interactive Technologies (CITI-USP), associated professor in the University of Vale do Rio do Sinos (Unisinos) and Faculdades Integradas de Taquara (Faccat), both from Brazil. His research interests are virtual and augmented reality, immersion and presence in VR, and user perception in immersive and interactive environments.



**Ana Grasielle Corrêa** is Ph.D. in Electrical Engineering at the Polytechnic School in the University of Sao Paulo (USP), Brazil. She is currently a researcher in the Interdisciplinary Center of Interactive Technologies (CITI-USP) and assistant professor in the Mackenzie University, Brazil. Her research interests are virtual and augmented reality, usability and interactive systems.



**Roseli de Deus Lopes** is Ph.D. in Electrical Engineering at the Polytechnic School in the University of Sao Paulo (USP), Brazil. She is currently Associated Professor 3 at the Department of Electronical Engineering and Digital Systems of the Polytechnic School at the University of Sao Paulo (EP-USP), Brazil. She is also vice-coordinator of the Interdisciplinary Center of Interactive Technologies (CITI-USP). Since 1988 she is a researcher at the Laboratorio de Sistemas Integraveis (LSI-USP), Brazil, where she works as principal investigator of the Interactive Electronic Media Group. She also coordinates scientific dissemination initiatives and projects aimed at identifying and developing talents in Science and Engineering. For example, since 2003 is the general coordinator of the Brazilian Fair of Sciences and Engineering (Febrace). Her research interests are computer graphics, digital image processing, techniques and devices for human-computer interactions, virtual and augmented reality.



**Marcelo Zuffo** is Ph.D. in Electrical Engineering at the Polytechnic School in the University of Sao Paulo (USP), Brazil. He is currently Full Professor at the Department of Electronical Engineering and Digital Systems of the Polytechnic School at the University of Sao Paulo (EP-USP), Brazil. Since 2011, he is the coordinator at the Interdisciplinary Center of Interactive Technologies (CITI-USP). Also, he is a researcher at the Laboratorio de Sistemas Integraveis (LSI-USP), Brazil, coordinating research and development projects in the interactive electronic media field. In 2001 developed the first system of fully immersive virtual reality in Brazil called Digital CAVE. His research interests are virtual reality, computer graphics and visualization and engineering of interactive media.

## Affiliations

**Eduardo Zilles Borba**<sup>1,2</sup> • **Ana Grasielle Corrêa**<sup>1</sup> • **Roseli de Deus Lopes**<sup>2</sup> • **Marcelo Zuffo**<sup>2</sup>

Ana Grasielle Corrêa  
agrasi@lsi.usp.br

Roseli de Deus Lopes  
roseli@lsi.usp.br

Marcelo Zuffo  
mkzuffo@lsi.usp.br

<sup>1</sup> Interdisciplinary Center in Interactive Technologies of University of Sao Paulo (CITI-USP), Av. Prof. Lucio Martins Rodrigues 436/4, Sao Paulo 05508-020, Brazil

<sup>2</sup> Department of Electrical Engineering and Digital Systems of the Polytechnic School, University of Sao Paulo (USP), Av. Prof. Luciano Gualberto 158/3, Sao Paulo 05508-020, Brazil