A virtual reality photography application to assess spatial memory

M.-Carmen Juan[a], Miguel Estevan[a], Magdalena Mendez-Lopez[b], Camino Fidalgo[b], Javier Lluch[a], and Roberto Vivo[a]

[a]Instituto Universitario de Automática e Informática Industrial, Universitat Politècnica de València, C/Camino de Vera, s/n, 46022 Valencia, Spain
[b]Departamento de Psicología y Sociología, Universidad de Zaragoza, Zaragoza, IIS Aragón, Spain

e-mails: mcarmen@dsic.upv.es; miesmo4@inf.upv.es; mmendez@unizar.es; alvarezcamino@unizar.es; jlluch@dsic.upv.es; rvivo@upv.es

Corresponding author: M.-Carmen Juan (mcarmen@dsic.upv.es)

Abstract
In this paper, the first virtual reality (VR) photography application that incorporates elements on panoramic images to assess short-term spatial memory is presented. It also allows the interaction with these elements. A study was conducted (N=50) to compare participants' performance outcomes and subjective experience when using a VR application with and without panoramic photography. The results show that both applications are effective in assessing short-term spatial memory. The results with panoramic photography include: 1) It was effective for short-term recall since, after using the application, the participants were able to verbally recall the objects and place them without significant differences with respect to the objects correctly placed; 2) The performance outcomes were independent of age and gender; 3) The perceived level of presence was directly related to experiencing less cybersickness when using the application; 4) The level of satisfaction was directly related to perceived enjoyment, concentration, usability, competence, calmness, and expertise; 5) The more the familiarity with VR applications, the less the perceived cybersickness. Finally, the application offers two main advantages: 1) The evaluators can customize the environment by adding as many elements as desired to the panoramic photograph; 2) It could be especially suitable for groups with reduced mobility.

Keywords: virtual reality photography, panoramic photography, 360º photograph, cinematic virtual reality, spatial memory, Oculus Quest

---

1. Introduction
The technique known as Virtual Reality (VR) photography (also known as cinematic VR) allows the interactive display of wide-angle panoramic photographs (usually 360º). A 360º photograph is a panoramic image with which interaction is possible. This type of photograph surrounds the point from which the photo is taken, so all of the details of the environment in which the photograph is taken can be observed. There are currently different devices for taking panoramic photographs (e.g., the Insta360 ONE R camera). These types of cameras usually have different wide-angle lenses, so they cover 360º and internally merge the photographs taken with the different lenses to finally obtain a 360º photograph.

VR photography has not been exploited for assessing short-term spatial memory. Short-term spatial memory is the ability of humans to remember spatial stimuli for short periods of time [1]. Spatial memory is used to memorize relevant information, for example, the route to find a place previously visited or where belongings were left [2]. Tools related to spatial memory can be used for assessment and training because spatial memory is strongly linked to personal autonomy [3]. For assessment, these tools can help to identify difficulties that hinder people's independence [4].

In this paper, a VR photography application to assess spatial memory is presented. The objective in this work was to develop and validate a VR photography application for the assessment of spatial short-term memory. For validation, the VR photography application presented in this work was compared with another VR application without panoramic photography in which the environments were modeled in 3D. The main hypothesis of this work was that the VR photography application would be effective for short-term recall.

The paper is organized as follows. Section 2 presents the state of VR photography and the spatial memory assessed by computer. Section 3 focuses on the design and development of the VR photography application.

It also briefly explains the hardware and software that was used in the development and validation. Section 4 details the characteristics of the participants in the study, the measurements, the configuration of the panoramic photograph used as environment, and the protocol followed. Section 5 presents the analyses and results. Section 6 discusses this work and results. Finally, Section 7 presents the conclusions.

2. Related Work

2.1. Virtual reality photography

Panoramic photography has mainly been used to create virtual tours of open spaces [5] or of closed spaces, such as museums or churches [6]. An example of a closed space is the Imperial Cathedral in Königslutter (Germany) [6] or the Yingxian Wooden Pagoda (China) [7]. An example of an open space is the city of Wolverhampton (UK) [5] (https://tourmkr.com/F1yxQEojqj/10786094p&145.32h&85.51t). Since 360° photos and videos can be embedded into a website, this type of content can be viewed just using a web browser. YouTube offers the ability to view 360° videos and have the experience of riding on a Roller Coaster, being attacked by a shark, or observing a pride of lions in their habitat. Similarly, Facebook provides the same functionality for photos and videos. Thus, users can produce and share content created with cameras that are capable of generating 360° content. Another possibility for viewing this type of content is to use VR headsets or viewers. In this case, experiences can be more immersive since they can be viewed in 3D and the user has the feeling of depth. These VR headsets or viewers range from Google Cardboard to headsets like Oculus Quest, which is the headset that was used in this work.

PeakLensVR [8] was presented with the idea of creating panoramic content for social networks. PeakLensVR is a mobile application for collecting mountain images and transforming them into panoramic images. The panoramic images can be later viewed with annotated information about the peaks. These images can be viewed with VR devices and can be shared on social networks.

Panoramic images can also be used for different purposes. For instance, panoramic photographs have been used as virtual elements that are mixed with a real environment in an Augmented Reality (AR) system for the treatment of acrophobia [9], [10]. Those works verified whether the panoramic photography used in an AR system was capable of evoking the sensation of height in users. They carried out a study in which 41 users, without fear of heights participated [9]. In their study, the users walked through the same environment, both physically and using AR. Their conclusions were that the panoramic photography experience was very high and that this technology could help in the treatment of acrophobia. Nevertheless, the users were able to distinguish the panoramic photography environments from the real ones.

Panoramic images provided by Google's Street View Image API have also been used for several purposes. For example, panoramic images (street view images) and computer vision methods were used to automate the characterization of built environments for neighborhood effects research [11]. Specifically, the authors used panoramic images from Google's Street View Image API for three cities in the United States (Charleston, Chicago, and Salt Lake City) and convolutional neural networks. They developed log Poisson regression models to estimate associations between the characteristics of built environments and the individual prevalence of diabetes and obesity in Salt Lake City. From their analyses, they concluded that neighborhood conditions can influence chronic disease outcomes.

To our knowledge, there are two works that are most related to this work. The first work is an experiment that was conducted to determine the effects produced in participants who navigated the National Mall area of Washington, DC, using Google Street View [12]. The study considered two conditions (a monitor and a head-mounted display). The authors did not find significant differences in the users' accuracy of direction to reference points or their sense of presence between the two conditions. However, according to the data collected in a questionnaire, the authors argued that since the participants' travel experience was based on teleportation between waypoints, the visual cues, such as trees and buildings, seemed to have a greater influence in determining the directions to reference points than the egocentric clues (i.e., first-person perspective). Moreover, the authors concluded that the users were generally dissatisfied with the discontinuous way of travel between waypoints in the panoramic environment. The second work [13] presented an application that used 360º images to assess the ability to recognize objects previously seen in a 360º video. The authors compared classic paper-and-pencil tests versus their VR application with 360º images and videos. Different objects were presented by a clinician in a 360º video by using the Oculus Go© headset. The clinician was in an office with 27 objects. The clinician moved around the office and presented 10 target objects (out of the 27). In a second phase, the participants looked around the 360° image of the office to find and name all ten target objects. The participants only looked around the environment, but did not have any type of interaction with the environment. Their results showed a significant correlation for the performance outcomes between the standard test and the VR application.

2.2.Computer-Assessed spatial memory
Spatial memory can be assessed using paper and pencil tests [14], [15]. The implementation of computerized tools that use virtual or augmented reality offers advantages over pencil and paper tests. Computerized tools facilitate the obtaining and storage of objective indicators of the individual's performance on the task to be carried out (e.g., reaction times for each part of the task, successes and failures, etc.). This storage facilitates its subsequent statistical analysis. For their part, VR and AR applications offer additional advantages in terms of the assessment and training of people in real environments (i.e., lower economic costs). For these reasons, VR and AR technologies are being exploited in different fields, including the study of spatial memory. VR has previously been used to assess spatial memory in humans [16]–[21]. The first VR applications for the assessment of spatial memory used a monitor to visualize the tasks, and the interaction was basic, without involving users' physical movements [22]–[24]. Later works have included users' physical movement to interact with VR environments [25]–[27].

AR has been used for memory assessment to a lesser extent [28]–[31]. In several AR applications, image targets distributed in the real world have been used [28], [29]. SLAM-based AR has also been used to place virtual elements on flat surfaces (horizontal or vertical) without the need to add additional elements to the scene, using mobile devices [30], [31] or HoloLens [32]. Mainly visual stimuli have been used, but auditory [33] or tactile [34] stimuli have also been studied to assess spatial memory.

Therefore, this work is the first VR photography application that adds elements into the panoramic images that are intended to be integrated into the images, and also includes interaction with those added objects. Moreover, this is the first time that this combination is used for the assessment of object-location memory.

3. Design and development of the VR photography application
This section describes the phases of the VR photography application and details the hardware and software used.

3.1. The VR photography application
The main objective of the application is for the participants to go through a virtual environment in which panoramic photography is used, during which they look at some objects and try to remember their location. The application was designed and implemented with three phases: 1) Tutorial; 2) Learning; and 3) Evaluation. First, the evaluator must configure the environment for the task in two phases. This configuration only has to be done once and can be used as many times as desired. In addition, as many different configurations as desired can be stored. These two phases are: 1) Selection of the panoramic photographs; and 2) Addition of the desired objects. These phases are described below in detail.

Selection of the panoramic photographs. The evaluator must select the number of different environments (spheres) that the user will use during the learning and evaluation phases. The evaluator can choose up to five different environments to use in the same study. In the application, five rooms of the same house were used as environments. However, the evaluators can choose environments that are unrelated, if they need them for their study.

Addition of the desired objects. The evaluator must configure the environments by adding the desired objects. The evaluator can choose up to thirty different 3D objects to place in the chosen environment/s.

The phases to be carried out by the user are the following:

The tutorial phase. The user becomes familiar with the functionality of the VR photography application by using the tutorial. This tutorial guides the user by indicating in writing what to do at all times (e.g., how to place an object, how to change objects, etc.).

The learning phase. During this phase, the user must explore the environment or environments (up to 5). The user can switch between environments (spheres) by using the X and Y buttons on the controller for the left hand. The objects to remember are highlighted by a box with four white arrows (Fig. 1). The pointed object is highlighted by a box with blue arrows (Fig. 1). When the user points to an object and the arrows are blue, she/he must select it in order to mark it as seen. After this selection, the object is shown with a semi-transparent dome (Fig. 2). The user cannot exit this phase until all of the objects are marked as seen. This ensures that the user has seen all of the objects that she/he must remember. The user does not have a maximum time to complete this phase. Fig. 1 and Fig. 2 show two examples of this phase.

The evaluation phase. This phase assesses the ability of the users to remember the location of the objects that are added to the panoramic photography and learned in the previous phase. In this phase, the users must place the objects in their correct locations. The objects appear in the central part of the image in a preset order, which is the same for all users. The user must take the object and place it in the position that she/he considers to be correct. The user does not have to place the objects in exactly the same location. The user has a margin of error when placing the objects in a certain position. This margin is fixed as a sphere with a 50-centimeter radius from the correct point. If the user places the object correctly the first time, no message appears and the object is shown with the semi-transparent dome (Fig. 3). Then, the next object is displayed in the center of the screen (Fig 3). If, on the other hand, the user has not placed the object correctly, a message is displayed to indicate this and the user has two more attempts. If the object is placed in an incorrect position on the third attempt, the object remains fixed in the position where it was placed, but the user is not informed that it was a failure in order to avoid an additional source of stress. Then, the next object appears in the center of the screen. The users do not receive any help regarding where to place the objects. Fig. 3 shows an example of this phase.

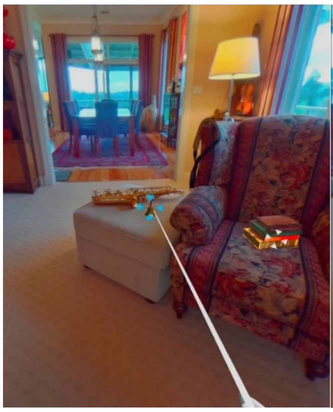

Fig. 1. Learning phase. Environment with two objects superimposed on the panoramic photograph: a bell highlighted with blue marks, and books highlighted with white marks.

Fig. 2. Learning phase. Environment with two objects superimposed on the panoramic photograph. The bell is covered with a semi-transparent dome to indicate that the user has seen it.

Fig 3. Evaluation phase. The user has already placed the bell, and the books appear in the center of the screen.

3.2. Development of the VR photography application (Hardware & Software)

A minimum requirement for the development and validation of the application was that the VR headset be a standalone device. Oculus Quest was selected because it is a standalone device, with no cables for either the controllers or the device. Since Oculus Quest integrates a processor and a storage unit, a computer is not needed to run the applications. The processor is Qualcomm Snapdragon 835, with 4 GB of RAM. It has two versions of storage, 64 GB and 128 GB. It has a Bluetooth module that allows both the controllers and any compatible device to be connected. It also has a Wi-Fi module. This headset incorporates four grayscale cameras on the front. Its function is to delimit the playing area in order to guarantee the user's safety. The playing area can be delimited by the user or automatically by the system, previously indicating the ground level. Moreover, these four cameras provide a hand recognition system to control applications. To do this, the system makes use of the images provided by the cameras to calculate the position of the hands in space. The only drawback is that the hands must always be within the viewing angle of the cameras. The device uses two wireless controllers (one for each hand), which communicate with the Bluetooth module of the headset. Each of these controllers has five buttons, two of which are triggers, and a joystick. For the lenses, each one of them has a resolution of 1440 x 1600 for each eye and a refresh rate of 72 Hz. The headset that is currently available for sale is Oculus Quest 2. The most notable differences compared to its predecessor are that Oculus Quest 2 has a resolution of 1832 x 1920 for each eye and a refresh rate of up to 90 Hz.

The VR photography application was developed using the Unity game engine (https://unity.com) and Oculus SDK (https://developer.oculus.com/downloads/package/unity-integration/) for the integration of Oculus in Unity. Thanks to the integration of Oculus in Unity, the compilation and installation of the application in the headset is fully managed by Unity. The Oculus SDK includes: 1) prefabs to manage the interaction with objects which can be a distance interaction or a direct interaction; 2) prefabs to manage the desired camera behavior; 3) prefabs to visualize the elements with which to interact with the environment. They can be the hands or the controllers; 4) prefabs to interact with the elements of the different interfaces; 5) an API to manage all of the events generated during the execution of the applications; 6) several examples.

Panoramic photographs have been used in Unity as textures applied to spheres. These textures must have the appropriate material shader (Skybox/Cubemap). The Oculus Quest camera (OVRCameraRig) must be inside the sphere with the panoramic photograph of the environment to be visited.

The objects overlapping the panoramic photograph are added inside the sphere so that they appear to be mixed with the objects of the panoramic photograph itself. All of the objects should be placed at the boundary of the sphere in order to avoid the levitation effect and to ensure the greatest possible realism.

When interacting with the added objects (e.g., picking up an object with the Oculus controllers), only one mesh is affected. In other words, if the object is made up of several meshes it is only interacting with one of these meshes. The Easy Mesh Combiner MT asset was used to combine the different meshes of an object (https://assetstore.unity.com/packages/tools/game-toolkits/easy-mesh-combiner-mt-138805).

4. Description of the study

This section describes the sample involved in the study, the measures used, the memory task, and the protocol followed.

4.1. Participants

A total of 50 participants were involved in the study. The study consisted of two groups, Group A and Group B. Group A performed the memory task using a VR photography application and Oculus Quest. This group consisted of 25 participants, ranging in age from 21 to 59 years old. The mean age was 26.12 ± 7.20 years old. There were 12 women (48%). Group B performed the memory task using a VR application without panoramic photography and Oculus Quest. This group consisted of 25 participants, ranging in age from 12 to 53 years old. The mean age was 26.28 ± 13.33 years old. There were 13 women (52%). Although there were six participants under the age of 18, they have not been removed from the sample because no significant differences were found when applying a Kruskal Wallis test to determine if age influences the performance variables. The participants or their parents were informed about the study and its objectives. They verbally accepted to participate in the study. The study was conducted in accordance with the declaration of Helsinki and was approved by the Ethics Committee of the Universitat Politècnica de València, Spain.

4.2.Measures

Performance variables with the memory task using the VR applications. The applications stored the following variables: the total number of objects located correctly during the evaluation phase (Total Objects); the total number of attempts made while placing the objects in the correct location (Attempts); and the total time required to complete the learning phase (Learning Time) and the evaluation phase (Evaluation Time).

Other performance variables. The participants had to verbally recall the objects they had seen in the memory task (variable: Objects Verbal). The participants had to place the objects they remembered on a printed map (variable: Objects Map).

Questionnaires. After completing the map-pointing task, the participants filled out a questionnaire about their subjective experience with the VR applications. The questionnaire consisted of 20 questions that were grouped into the following variables: no-cybersickness, enjoyment, concentration, usability, competence, calmness, no-mental effort, no-physical effort, satisfaction, ergonomics, and presence. The questionnaire also included five more questions related to their familiarity with computers, videogames, or VR applications and two open-ended questions. This questionnaire was designed specifically for this study, and some of the questions were adapted from commonly used questionnaires [35]–[37] and based on our previous experiences [31]. The participants filled out the questionnaire online using a web browser. All of the questions were formulated in a positive way. All of the questions were on a 7-point Likert scale, ranging from 1 "Totally disagree" to 7 "Totally agree". Scores for the variables were obtained by calculating the mean value of the associated questions.

4.3. Memory task

The memory task used a panoramic photograph of a room (https://eliteagent.com/360-rooms-app-allows-creation-virtual-tours-properties). The room had some furniture that can be found in a dining room. The objects that were added were objects that could be found in that type of environment. Eight objects were used in this study. The eight objects used in the study were: a plant, a brush, a sculpture, a slipper, a cup and saucer, a toy car, a bell, and some books (Fig. 4). The shape of the room and the location of the eight objects is shown in Fig. 5. Fig. 6 has two views of the environment showing two objects that are integrated in each of the two views.

Fig. 4. The eight objects used in this study.

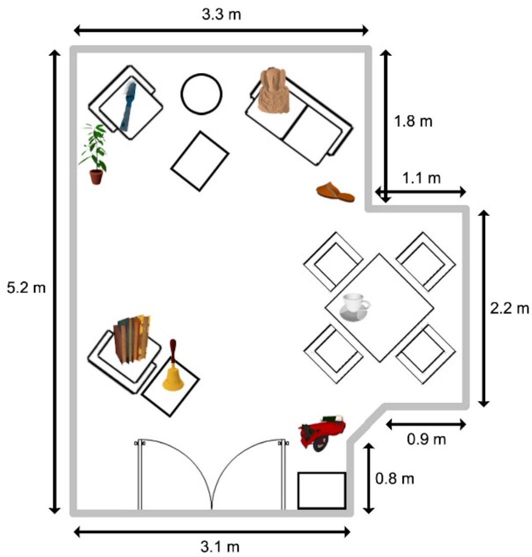

Fig. 5. Shape of the room and location of the eight objects used in this study. The eight objects are highlighted with red circles/ovals.

Fig. 6. Two views of the environment in which the two objects integrated in each of the two views can be seen.

Fig. 7. A user with the Oculus Quest and controllers.

4.4. Procedure

The participants were counterbalanced and randomly assigned to one of two conditions:

- Group A (PanoGroup): The participants who performed the memory task using the VR photography application using Oculus Quest.
- Group B (NoPanoGroup): The participants who performed the memory task using the VR application with no panoramic photography using Oculus Quest.

The protocol was the following:

1. The participants performed the memory task (PanoGroup or NoPanoGroup).
2. The participants performed the object-recall task, which consisted of free recall of the eight objects that were learned using the two VR applications. The supervisor asked the participants the following question: "Which objects do you remember seeing when using the VR application?" The participants verbally indicated the objects and the supervisor wrote them down. No feedback was given on responses.
3. The participants performed the map-pointing task, which tested the ability of the participants to read a bi-dimensional map of the room from their memory of the virtual room. The participants had to place each object they remembered on a printed map in its correct location according to the memory task. The participants had a sheet showing the objects with an assigned letter. The participants only had to write the letters on the area of the map.
4. The participants filled out an online questionnaire using a web browser.

5. Results

This section details the statistical analysis conducted with the data obtained during this study. An initial descriptive analysis was carried out to explore different measures (e. g., means, standard deviations). The Shapiro-Wilk test was used to check the normal distribution of the variables. The tests indicated that the sample did not fit a normal distribution. Accordingly, non-parametric tests were used (the Mann-Whitney U test, the Wilcoxon signed-rang test, the Kruskal-Wallis test, and the Spearman correlation). The results were considered to be statistically significant if $p < 0.05$. The statistical open source toolkit R (http://www.r-project.org) was used to carry out the statistical analysis of the data. Fig. 7 shows a user with the Oculus Quest and controllers.

5.1. Performance outcomes

In order to determine how the use of panoramic photography affects the performance outcomes of the participants, the performance outcomes between the two groups (the PanoGroup vs. the NoPanoGroup) were compared. Fig. 8 shows box plots for the performance outcome variables and for the two groups (PanoGroup and NoPanoGroup).

To determine whether or not there were differences in the performance variables between the participants of the two groups, the Mann-Whitney U tests were applied. The results are shown in Table 1. Statistical differences were found for the performance variables except for the Total Objects variable. These results can also be observed in Fig. 8. If the outliers are eliminated and the Mann-Whitney U tests are applied again, significant differences remain for the same three performance variables.

The Wilcoxon Signed-rank test was applied to determine whether or not there were differences between the total number of objects correctly placed using the VR photography application and the total number of objects that the participants remember and verbally indicate after performing the task (Objects Verbal), and differences between the total number of objects correctly placed using the VR photography application and the total number of objects correctly placed in the pointing-map task (Objects Map). The results are shown in Table 2. Table 2 shows no statistically significant differences for the recall variables.

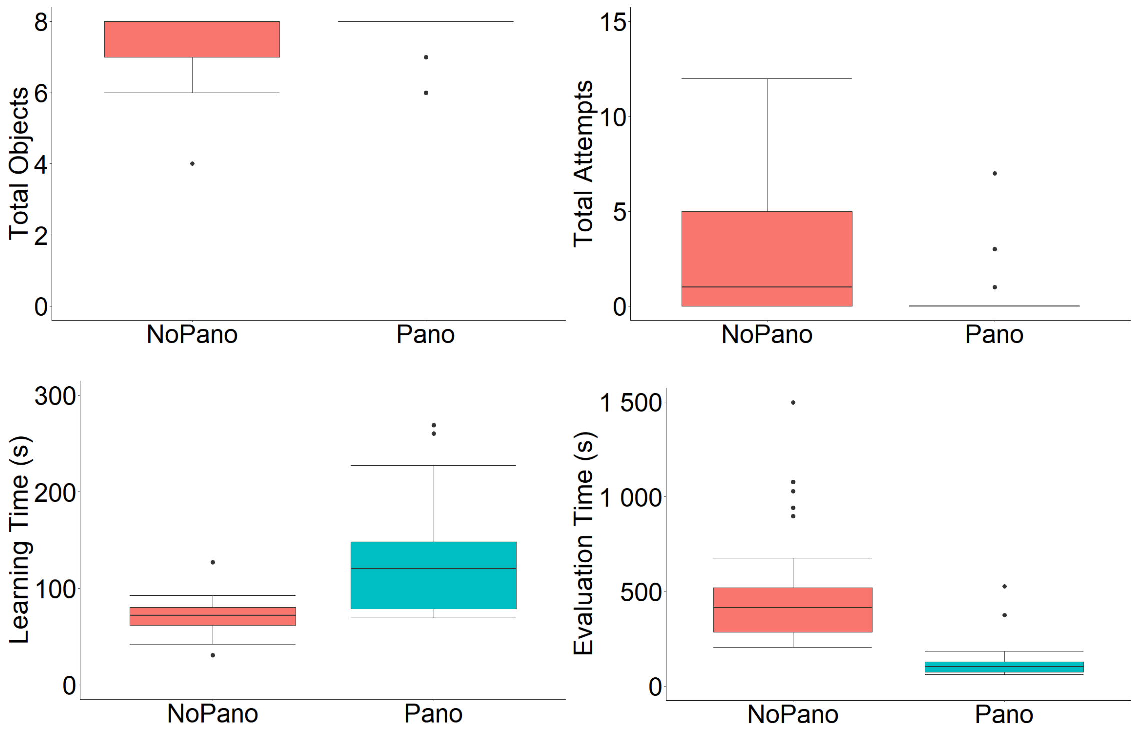

Fig. 8. Box plots for the performance outcome variables and for the two groups (panoramic photography (Pano) vs. no panoramic photography (NoPano)).

Table 1. Mann-Whitney U test for the performance variables and considering the two groups: Panoramic Photography (P) vs. No-Panoramic Photography (NoP)

	P	*NoP*	*U*	*Z*	*p*	*r*
Total Objects	8;0	8;1	378.5	1.914	0.058	0.271
Attempts	0;0	1;5	179.5	-2.885	**0.004**	0.408
Learning Time	120;69	71;18	522.0	4.065	**0.001**	0.575
Evaluation Time	101;53	412;234	29.0	-5.501	**0.001**	0.778

Note: The values in columns P and NoP depict the median and the interquartile range, respectively (e.g., 8;0). The numbers in bold indicate significant differences.

Table 2. Wilcoxon signed-rank test for the recall variables: the total number of objects correctly placed using the VR photography application (vrp), the total number of objects verbally remembered (verbal), and the total number of objects correctly placed in the pointing-map task (map)

VRP	*VERBAL*	*MAP*	*U*	*Z*	*p*	*r*
8;0	8;0		1.5	-0.601	0.586	0.085
8;0		8;0	11.5	1.320	0.341	0.187
	8;0	8;0	10.0	1.998	0.098	0.283

The values in columns VRP, VERBAL and MAP depict the median and the interquartile range, respectively. The variables are compared two by two. The two variables compared are those that include the median and the interquartile range.

5.2. Gender and age comparisons

The Mann Whitney U tests were applied to check if gender affected the performance outcome variables. For these analyses, the data of two participants who were outliers in all of the performance variables considered were eliminated from the sample. The results are shown in Table 3. Fig. 9 shows the interaction plots for the performance variables, considering the age and gender of the participants of the sample analyzed in this section. As can be observed in Table 3, no statistically significant differences were found for any of the performance variables. From these results, it can be concluded that the performance results were independent of the gender of the participants when using the VR application with panoramic photography.

The Kruskal Wallis test was applied to determine if age influences the performance variables. For these analyses, the data of two participants who were outliers in all of the performance variables considered were eliminated from the sample. The results of this analysis are shown in Table 4. From these analyses, it can be concluded that the performance results were independent of the age of the participants when using the VR photography application.

Table 3. Mann-Whitney U test for the performance variables, considering gender (women vs. men), when using the VR photography application and Oculus Quest

	Women	*Men*	*U*	*Z*	*p*	*r*
Total Objects	8;0	8;0	58.5	-1.140	0.293	0.238
Attempts	0;0.75	0;0	75.5	0.906	0.388	0.189
Learning Time	120;79	101;52	62.0	-0.186	0.879	0.039
Evaluation T.	114;73	78;29	92.0	1.674	0.101	0.349

Note: The values in columns Women and Men depict the median and the interquartile range, respectively.

Table 4. Kruskal Wallis test for the performance variables, considering the age of the participants, when using the VR photography application and Oculus Quest

	d.f.	*H*	*p*
Total Objects	8	10.50	0.232
Attempts	8	13.94	0.083
Learning Time	8	7.79	0.454
Evaluation Time	8	8.35	0.400

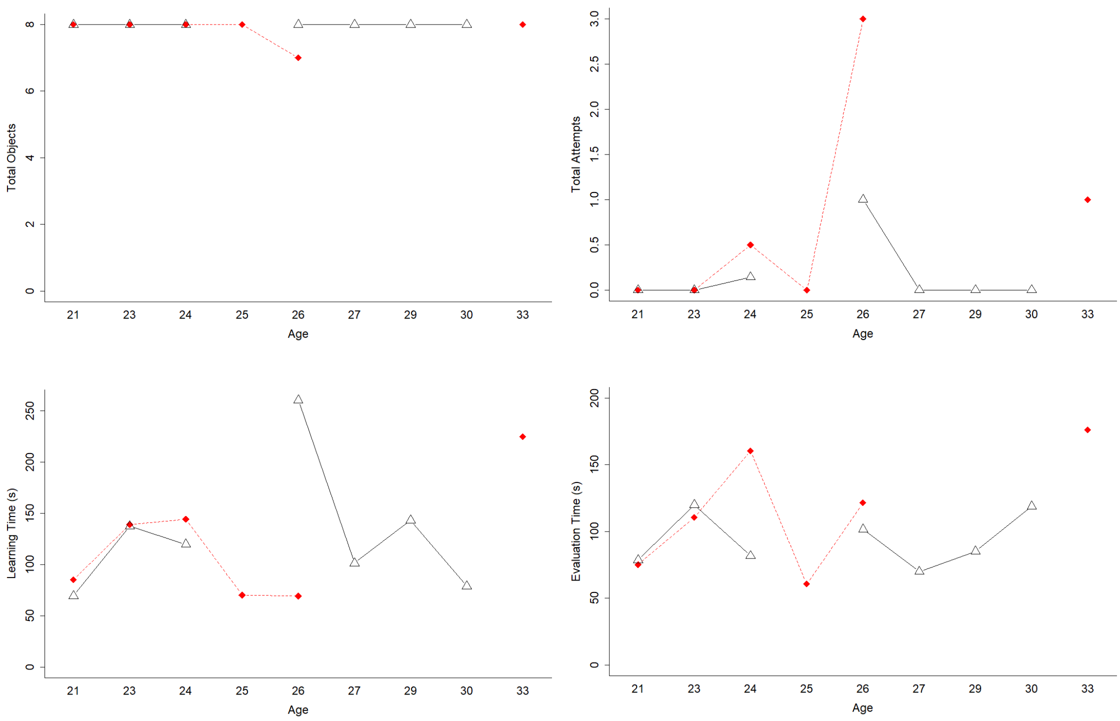

Fig. 9. Interaction plots for the performance outcome variables for the two groups, considering the age and gender of the participants. The red rhombuses indicate women.

5.3. Subjective variables

The Mann-Whitney U tests were applied to check whether or not there were differences for the subjective variables between the participants of the two groups (PanoGroup vs. NoPanoGroup). Table 5 shows the results. The results indicate that there is only a statistically significant difference for the No cybersickness variable in favor of the participants in Group B (No panoramic photography), who felt significantly less dizziness during or after the experience. Fig. 10 shows a radial graph with the means of the different subjective variables from the questionnaire filled out by the users after using the two VR applications and Oculus Quest. The difference with respect to the No cybersickness variable can be observed in Fig. 10.

The Mann-Whitney U tests were applied to check whether or not there were differences for the subjective variables after using the VR photography application and considering the gender. For these analyses, the data of two participants who were outliers in all of the performance variables considered were eliminated from the sample. Table 6 shows the results.

Table 5. Mann-Whitney U test for the subjective variables and considering the two groups: Panoramic Photography (P) vs. No-Panoramic Photography (NoP)

	P	*NoP*	*U*	*Z*	*p*	*r*
No Cybersick	7;2.5	7;0	232.0	-2.156	**0.032**	0.305
Enjoyment	7;0	7;1	330.0	0.441	0.668	0.062
Concentration	6.5;2	6.5;1.5	306.0	-0.130	0.904	0.018
Usability	6.3;1	6.3;1.3	281.5	-0.613	0.547	0.087
Competence	7;1	7;1	345.0	0.769	0.449	0.109
Calmness	7;1	7;1	343.5	0.731	0.472	0.103
Expertise	6;2	6;2	280.5	-0.646	0.525	0.091
No Mental E	6;0	7;1	226.0	-1.796	0.074	0.254
No Physical E	7;0	7;0	327.0	0.541	0.602	0.076
Ergonomics	6;1	7;1	275.0	-0.802	0.429	0.113
Satisfaction	6.5;1	7;0.6	223.0	-1.831	0.069	0.259
Presence	5.5;1.75	6.25;2.5	286.0	-0.518	0.611	0.073

Note: The values in columns P and NoP depict the median and the interquartile range, respectively. The numbers in bold indicate significant differences.

Table 6. Mann-Whitney U test for the subjective variables after using the Panoramic Photography application, considering the gender (women vs. men)

	Women	*Men*	*U*	*Z*	*p*	*r*
No Cybersick	6.25;2.9	7;0	45	-1.459	0.155	0.304
Enjoyment	7;0	7;0	73	0.751	0.481	0.157
Concentration	6.25;2.38	6.5;1	53	-0.773	0.459	0.161
Usability	6.33;1.17	6.67;1	54	-0.694	0.508	0.145
Competence	7;0	7;0	67	0.173	0.896	0.036
Calmness	7;0.75	7;0	62	-0.243	0.840	0.051
Expertise	6;1	6;1	42	-1.501	0.142	0.313
No Mental E	6;0.75	6;0	70	0.319	0.776	0.067
No Physical E	7;0	7;0	52	-1.650	0.113	0.344
Ergonomics	6;0.75	7;1	42	-1.565	0.126	0.326
Satisfaction	6.5;1.63	6.5;1	53	-0.808	0.438	0.168
Presence	5.88;1.94	5.25;1	81	0.967	0.349	0.202

Note: The values in columns Women and Men depict the median and the interquartile range, respectively.

5.4. Correlations

Table 7 shows the correlations obtained between the subjective variables when using the VR photography application and Oculus Quest. This table shows that the perceived level of presence is directly related to experiencing fewer adverse effects when using the application. As an example, this correlation is shown graphically in Fig. 11. Perceived expertise is directly related to experiencing fewer adverse effects, perceived enjoyment, concentration, usability, competence, and calmness. The level of satisfaction is directly related to perceived enjoyment, concentration, usability, competence, calmness, and expertise.

The data that the participants provided regarding their level of experience with computers, experience with video games, regularity as players or familiarity with VR were analyzed. Table 8 shows the correlations between the subjective variables and considering the participants of the group that used panoramic photography. From these results, it can be highlighted that, the more experience in computing, the more competence and expertise perceived, but the lower the level of presence experienced. The more video game experience, the greater the usability, and the more perceived competence and expertise. The more familiarity with VR applications, the fewer perceived adverse effects and the higher the level of expertise.

Table 7. Spearman correlation. correlations between subjective variables, considering the VR photography application and Oculus Quest

	2	3	4	5	6	7	8	9	10
1	0.28	0.17	0.38	0.36	0.18	**0.47**	0.35	0.09	**0.57**
2	1.00	**0.44**	0.39	**0.47**	**0.54**	**0.43**	0.09	**0.40**	0.34
3		1.00	**0.71**	**0.58**	0.39	**0.69**	**0.44**	**0.82**	-0.03
4			1.00	**0.70**	**0.42**	**0.80**	**0.54**	**0.52**	0.17
5				1.00	0.24	**0.65**	0.35	**0.45**	0.00
6					1.00	**0.50**	0.29	**0.44**	0.31
7						1.00	**0.45**	**0.65**	0.19
8							1.00	0.31	0.15
9								1.00	0.17
10									1.00

Note: The meaning of the numbers in the first row and in the first column is as follows: 1- No cybersickness; 2- Enjoyment; 3- Concentration; 4- Usability; 5- Competence; 6- Calmness; 7- Expertise; 8- Ergonomics; 9- Satisfaction; 10- Presence. Numbers in bold indicate significant correlations.

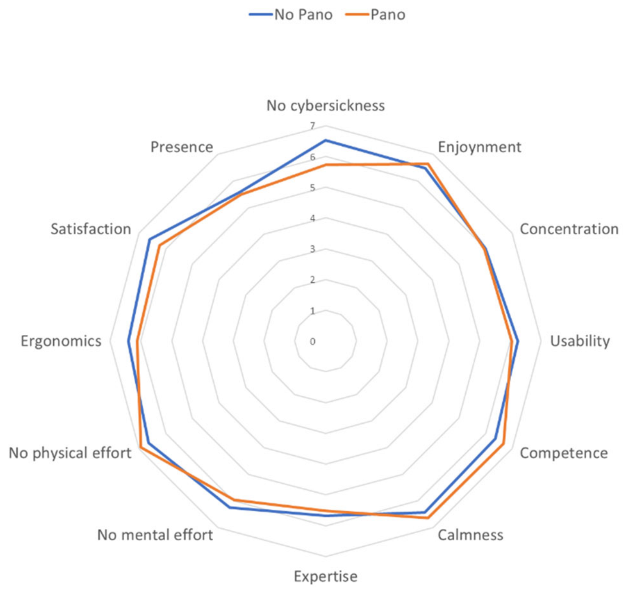

Fig. 10. Radial graph showing the mean scores for the subjective variables, considering the two groups.

Table 8. Spearman correlation. correlations between the variables of the participants' experience with technology and subjective variables, considering the VR photography application and Oculus Quest

	2	3	4	5	6	7	8	9
1	**0.59**	**0.54**	**0.51**	0.10	0.31	**0.43**	**0.41**	**-0.47**
2	1.00	**0.84**	**0.70**	0.33	**0.46**	**0.44**	**0.53**	-0.11
3		1.00	**0.58**	0.29	0.32	**0.49**	**0.43**	-0.17
4			1.00	**0.54**	0.28	0.22	**0.44**	0.08

Note: The meaning of the numbers in the first row and in the first column is as follows: 1- Experience with computers; 2- Experience with video games; 3- Regularity as a player; 4- familiar with RV; 5- No cybersickness; 6- Usability; 7- Competence; 8- Expertise; 9- Presence. The numbers in bold indicate significant correlations. Only the subjective variables with which there was any significant correlation were included.

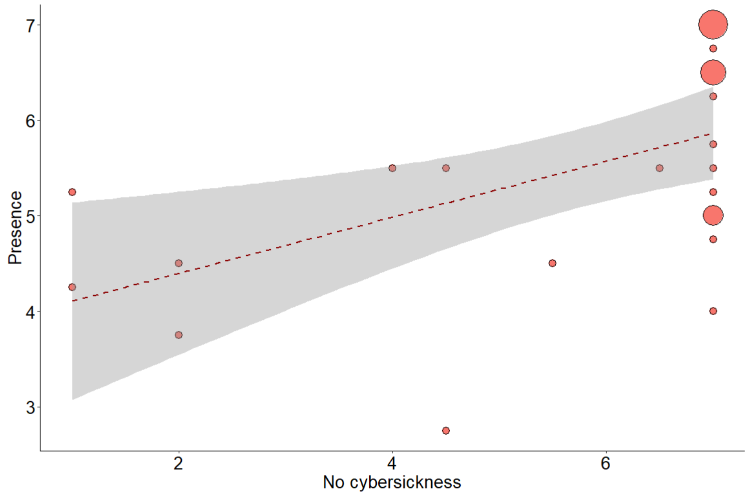

Fig. 11. Scatter plots for the significant correlation between the No cybersickness and presence variables. Four sizes of circles appear in the plot that represent the number of occurrences. The dashed red lines are the best-fitting linear regression lines. The grey area represents a 95% confidence level interval for predictions from a linear model.

6. Discussion

This is the first work that combines panoramic photography with the inclusion of elements that can be integrated into the photograph and appear to be part of the photograph itself, in addition to allowing interaction with them. From our point of view, this fusion offers many possibilities, other than the assessment of spatial memory that can be exploited for other purposes.

The Oculus Quest was used as a headset, which is a standalone device that does not have to be physically connected to other devices. Therefore, it allows freedom of movement.

From the results, there was no significant difference in the number of objects placed correctly between the two VR applications (Pano vs. NoPano). However, the participants in the PanoGroup required significantly less time to complete the evaluation phase and also required fewer attempts. An explanation for these results is that, in the PanoGroup, the participants only have to turn the chair to see a different part of the environment. In the NoPanoGroup, the participants have to physically walk through the real environment in order to place the objects. This walking involves a higher cognitive workload, which makes the task more difficult (requiring more attempts) and more time consuming. In addition, in the PanoGroup, the environment provides the participants with several environmental clues for easy viewing. Also, the egocentric clues in this condition are very basic (i.e., turns while sitting in a chair) in comparison with the NoPanoGroup (i.e., path followed while walking). With regard to the learning time, in the NoPanoGroup, the participants can see all of the objects at a glance without needing to turn around to see the entire environment. Therefore, both applications have been shown to be effective in assessing short-term spatial memory. Each application has different features that make it more suitable for certain groups. The VR photography application could be especially suitable for groups with reduced mobility.

The results indicate that the VR photography application is suitable for the assessment of spatial memory regardless of the gender of the participants. This result is in line with previous works [28], [34]. The VR photography application is also suitable for the assessment of spatial memory regardless of the age of the participants. This result is also in line with previous works [28].

With regard to the recall variables (Total Objects, Objects Map, and Objects Verbal), no statistically significant differences were found. These results show that spatial-visual associations were learned and transferred from the three-dimensional array of the real room to the bi-dimensional array of the map and the mental image of the environment. This result corroborates the main hypothesis (the VR photography application would be effective for short-term recall).

With regard to the learning effectiveness of this proposal and previous related works, to our knowledge, there are only two works that have used panoramic photography for spatial memory [12], [13] introduced in Section 2.1. In [12], an experiment was conducted to determine the effects produced in participants who navigated the National Mall area of Washington, DC, using Google Street View. The study considered two conditions (a monitor and a head-mounted display). The authors did not find significant differences between the two conditions in users' accuracy of direction to reference points. In that work, the panoramic images were used in both conditions and the participants had to travel between waypoints in the panoramic environment. In any case, the panoramic images were effective in learning the path to a destination target. In [13], the results showed a significant correlation for the performance outcomes between the standard test and the VR application to assess the ability to recognize objects previously seen. That effectiveness in is line with the results obtained in this work.

With regard to the subjective variables, when comparing the two VR applications, the results indicated that there was only a statistically significant difference for the No cybersickness variable. The participants who used the VR application with no panoramic photography felt significantly less dizziness during or after the experience. However, if only the gender of the participants that used the VR photography application is considered, the results indicated that the subjective experience of the participants was independent of gender.

If correlations are analyzed for the VR photography application, the perceived level of presence is directly related to experiencing fewer adverse effects when using the application. Another important conclusion is that the level of satisfaction is directly related to perceived enjoyment, concentration, usability, competence, calmness, and expertise. Moreover, the more the familiarity with VR applications, the fewer the perceived adverse effects.

This proposal is the first VR photography application that incorporates elements into panoramic images for the assessment of spatial memory, but it is not the first time that VR [25], [27], [38], [39] or AR [28], [29], [32], [34] has been used for the assessment of working memory or spatial memory. There are substantial differences in the studies previously carried out and this work. For instance, in the study of Rodríguez-Andrés et al. [25], the virtual environment was an outdoor park in which a series of tables were distributed on which the objects that users had to remember appeared. They used the same environment and compared two types of interactions (standard interaction vs. natural interaction involving the user's physical movement). The VR system was effective in learning and was independent of gender. These results are in line with the results obtained in this work. In the work of Wais et al. [38], they developed a VR labyrinth, and the participants used the HTC VIVE as a headset, which is not a standalone device. They carried out a study comparing two groups (the Labyrinth-VR vs. placebo control games). Their results suggested an improvement in high-fidelity, long-term memory capability for the Labyrinth-VR group. In this work, short-term memory was considered, not long-term memory. A possible future study would be to use panoramic photography to determine its advantages for long-term memory. In the study of Sin et al. [39], they developed two totally different VR environments (underwater and Mars). They used a wireless HTC Vive Pro headset to study how the effect of memory depends on the context. The work of Pieri et al. [13] is more aligned with this work. They presented an application that used 360º images to assess the ability to recognize objects previously seen in a 360º video. However, they do not assess object-location memory, add objects to the 360º content, or interact with added objects, and the study was also different. Very diverse environments can be created using panoramic photography; to do so, only panoramic content would have to be available.

Finally, it can be highlighted that both videos and images with panoramic content can be found and used for different purposes and studies. For example, Google Street View images are a very valuable resource (e.g., [11]).

7. Conclusions

The first VR photography application that incorporates virtual elements (and the interactions with these elements) in panoramic photography was presented. With this new proposal, an application was developed and validated to assess short-term spatial memory. This study was carried out using Oculus Quest as a headset. As a standalone device, Oculus Quest allows freedom of movement without having to be physically connected to any other device. Since the application uses panoramic photography, the environment can be created as desired. The evaluator only has to obtain a panoramic photograph of the environment or use a panoramic photograph that meets the needs of a specific study.

For the first time, the results of using an environment with a panoramic photograph and using a 3D modeled environment were compared. From the results, it can be concluded that both types of applications can be used to assess spatial memory. The use of panoramic photography was effective for short-term recall since, after using the application, the participants were able to verbally recall the objects and place them without significant differences when compared to the objects correctly placed using the VR application. The performance outcomes using the VR photography application were independent of age and gender. However, more research is needed to further investigate how panoramic photography can help to assess spatial memory and identify other uses, not just spatial memory.

Acknowledgments

We would like to thank Elite Agent as well as Sergio Alcaraz, Ramon Mollá, and all of the people who participated in the study.

Sources of support:

MCIN/ AEI /10.13039/501100011033/ and by "ERDF A way of making Europe" through the project AR3Senses (TIN2017-87044-R); Gobierno de Aragón (research group S31_20D) and FEDER 2020-2022 "Construyendo Europa desde Aragón"

References

[1] A. Baddeley, "Working memory," *Science (80-.).*, vol. 255, no. 5044, pp. 556–559, 1992, doi: 10.1126/science.1736359.

[2] N. Burgess, S. Becker, J. A. King, and J. O'Keefe, "Memory for events and their spatial context: Models and experiments," *Philos. Trans. R. Soc. B Biol. Sci.*, vol. 356, no. 1413, pp. 1493–1503, 2001, doi: 10.1098/rstb.2001.0948.

[3] L. A. Cushman, K. Stein, and C. J. Duffy, "Detecting navigational deficits in cognitive aging and Alzheimer disease using virtual reality.," *Neurology*, vol. 71, no. 12, pp. 888–895, Sep. 2008, doi: 10.1212/01.wnl.0000326262.67613.fe.

[4] A. Neguţ, S. A. Matu, F. A. Sava, and D. David, "Task difficulty of virtual reality-based assessment tools compared to classical paper-and-pencil or computerized measures: A meta-analytic approach," *Comput. Human Behav.*, vol. 54, pp. 414–424, 2016, doi: 10.1016/j.chb.2015.08.029.

[5] E. Ramsey, "Virtual Wolverhampton: Recreating the historic city in virtual reality," *Archnet-IJAR*, vol. 11, no. 3, pp. 42–57, 2017, doi: 10.26687/archnet-ijar.v11i3.1395.

[6] A. P. Walmsley and T. P. Kersten, "The imperial cathedral in Königslutter (Germany) as an immersive experience in virtual reality with integrated 360° panoramic photography," *Appl. Sci.*, vol. 10, no. 4, 2020, doi: 10.3390/app10041517.

[7] D. S. Lai, A. K. H. Leung, D. Chan, and S. H. Ching, "Cultural heritage preservation using new media methods: Yingxian Wooden Pagoda, China," *Virtual Archaeol. Rev.*, vol. 10, no. 21, pp. 103–115, 2019, doi: 10.4995/var.2019.11071.

[8] A. La Salandra, D. Frajberg, and P. Fraternali, "A virtual reality application for augmented panoramic mountain images," *Virtual Real.*, vol. 24, no. 1, pp. 123–141, 2020, doi: 10.1007/s10055-019-00385-x.

[9] M. C. Juan, R. Baños, C. Botella, D. Pérez, M. Alcaniz, and C. Monserrat, "An Augmented Reality System for the Treatment of Acrophobia: The Sense of Presence Using Immersive Photography," *Presence Teleoperators Virtual Environ.*, vol. 15, no. 4, pp. 393–402, Aug. 2006, doi: 10.1162/pres.15.4.393.

[10] M. C. Juan *et al.*, "An augmented reality system for the treatment of acrophobia," in *The 8th International Workshop on Presence (PRESENCE 2005)*, 2005, pp. 315–318.

[11] Q. C. Nguyen *et al.*, "Neighbourhood looking glass: 360 automated characterisation of the built environment for neighbourhood effects research," *J. Epidemiol. Community Health*, vol. 72, no. 3, pp. 260–266, 2018, doi: 10.1136/jech-2017-209456.

[12] P. E. Napieralski *et al.*, "An evaluation of immersive viewing on spatial knowledge acquisition in spherical panoramic environments," *Virtual Real.*, vol. 18, no. 3, pp. 189–201, 2014, doi: 10.1007/s10055-014-0245-1.

[13] L. Pieri, S. Serino, P. Cipresso, V. Mancuso, G. Riva, and E. Pedroli, "The ObReco-360°: a new ecological tool to memory assessment using 360° immersive technology," *Virtual Real.*, 2021, doi: 10.1007/s10055-021-00526-1.

[14] J. Langlois, C. Bellemare, J. Toulouse, and G. A. Wells, "Spatial abilities and technical skills performance in health care: A systematic review," *Med. Educ.*, vol. 49, pp. 1065–1085, 2015, doi: 10.1111/medu.12786.

[15] M. Mitolo, S. Gardini, P. Caffarra, L. Ronconi, A. Venneri, and F. Pazzaglia, "Relationship between spatial ability, visuospatial working memory and self-assessed spatial orientation ability: a study in older adults," *Cogn. Process.*, vol. 16, no. 2, pp. 165–176, 2015, doi: 10.1007/s10339-015-0647-3.

[16] C. J. Bohil, B. Alicea, and F. A. Biocca, "Virtual reality in neuroscience research and therapy," *Nat. Rev. Neurosci.*, vol. 12, no. 12, pp. 752–762, 2011, doi: 10.1038/nrn3122.

[17] I. León, L. Tascón, and J. M. Cimadevilla, "Age and gender-related differences in a spatial memory task in humans," *Behav. Brain Res.*, vol. 306, pp. 8–12, 2016, doi: 10.1016/j.bbr.2016.03.008.

[18] S. Münzer and M. V. Zadeh, "Acquisition of spatial knowledge through self-directed interaction with a virtual model of a multi-level building: Effects of training and individual differences," *Comput. Human Behav.*, vol. 64, pp. 191–205, 2016, doi: 10.1016/j.chb.2016.06.047.

[19] I. León, L. Tascón, J. J. Ortells-Pareja, and J. M. Cimadevilla, "Virtual reality assessment of

walking and non-walking space in men and women with virtual reality-based tasks," *PLoS One*, vol. 13, no. 10, pp. 1–15, 2018, doi: 10.1371/journal.pone.0204995.
[20] H. Fabroyir and W. C. Teng, "Navigation in virtual environments using head-mounted displays: Allocentric vs. egocentric behaviors," *Comput. Human Behav.*, vol. 80, pp. 331–343, 2018, doi: 10.1016/j.chb.2017.11.033.
[21] M. Jonson, S. Avramescu, D. Chen, and F. Alam, "The Role of Virtual Reality in Screening, Diagnosing, and Rehabilitating Spatial Memory Deficits," *Front. Hum. Neurosci.*, vol. 15, p. 32, 2021, doi: 10.3389/fnhum.2021.628818.
[22] S. Walkowiak, T. Foulsham, and A. F. Eardley, "Individual differences and personality correlates of navigational performance in the virtual route learning task," *Comput. Human Behav.*, vol. 45, pp. 402–410, 2015, doi: 10.1016/j.chb.2014.12.041.
[23] L. Picucci, A. O. Caffò, and A. Bosco, "Besides navigation accuracy: Gender differences in strategy selection and level of spatial confidence," *J. Environ. Psychol.*, vol. 31, no. 4, pp. 430–438, 2011, doi: 10.1016/j.jenvp.2011.01.005.
[24] J. M. Cimadevilla, J. R. Lizana, M. D. Roldán, R. Cánovas, and E. Rodríguez, "Spatial memory alterations in children with epilepsy of genetic origin or unknown cause," *Epileptic Disord.*, vol. 16, no. 2, pp. 203–207, 2014, doi: 10.1684/epd.2014.0661.
[25] D. Rodríguez-Andrés, M.-C. Juan, M. Méndez-López, E. Pérez-Hernández, and J. Lluch, "MnemoCity Task: Assessment of Childrens Spatial Memory Using Stereoscopy and Virtual Environments," *PLoS One*, vol. 11, no. 8, p. e0161858, Aug. 2016, doi: 10.1371/journal.pone.0161858.
[26] D. Rodriguez-Andres, M. Mendez-Lopez, M.-C. Juan, and E. Perez-Hernandez, "A Virtual Object-Location Task for Children: Gender and Videogame Experience Influence Navigation; Age Impacts Memory and Completion Time," *Front. Psychol.*, vol. 9, p. 451, Apr. 2018, doi: 10.3389/fpsyg.2018.00451.
[27] S. Cárdenas-Delgado, M. Méndez-López, M. C. Juan, E. Pérez-Hernández, J. Lluch, and R. Vivó, "Using a virtual maze task to assess spatial short-term memory in adults," in *VISIGRAPP 2017 - Proceedings of the 12th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications*, 2017, pp. 46–57.
[28] M.-C. Juan, M. Mendez-Lopez, E. Perez-Hernandez, and S. Albiol-Perez, "Augmented reality for the assessment of children's spatial memory in real settings," *PLoS One*, vol. 9, no. 12, p. e113751, Dec. 2014, doi: 10.1371/journal.pone.0113751.
[29] M. Mendez-Lopez, E. Perez-Hernandez, and M.-C. Juan, "Learning in the navigational space: Age differences in a short-term memory for objects task," *Learn. Individ. Differ.*, vol. 50, pp. 11–22, Aug. 2016, doi: 10.1016/j.lindif.2016.06.028.
[30] F. Munoz-Montoya, C. Fidalgo, M.-C. Juan, and M. Mendez-Lopez, "Memory for Object Location in Augmented Reality: The Role of Gender and the Relationship Among Spatial and Anxiety Outcomes," *Front. Hum. Neurosci.*, vol. 13, p. 113, Mar. 2019, doi: 10.3389/fnhum.2019.00113.
[31] F. Munoz-Montoya, M.-C. Juan, M. Mendez-Lopez, and C. Fidalgo, "Augmented Reality Based on SLAM to Assess Spatial Short-Term Memory," *IEEE Access*, vol. 7, pp. 2453–2466, 2019, doi: 10.1109/ACCESS.2018.2886627.
[32] J. Keil, A. Korte, A. Ratmer, D. Edler, and F. Dickmann, "Augmented Reality (AR) and Spatial Cognition: Effects of Holographic Grids on Distance Estimation and Location Memory in a 3D Indoor Scenario," *PFG - J. Photogramm. Remote Sens. Geoinf. Sci.*, vol. 88, pp. 165–172, 2020, doi: 10.1007/s41064-020-00104-1.
[33] M. Loachamín, M.-C. Juan, M. Mendez-Lopez, E. Pérez-Hernández, and M. J. Vicent, "Developing and Evaluating a Game for the Assessment of Spatial Memory Using Auditory Stimuli," *IEEE Lat. Am. Trans.*, vol. 13, no. 10, pp. 1653–1661, 2019.
[34] F. Munoz-Montoya, M. C. Juan, M. Mendez-Lopez, R. Molla, F. Abad, and C. Fidalgo, "SLAM-based augmented reality for the assessment of short-Term spatial memory. A comparative study of visual versus tactile stimuli," *PLoS One*, vol. 16, no. 2 February, pp. 1–30, 2021, doi: 10.1371/journal.pone.0245976.
[35] B. G. Witmer and M. J. Singer, "Measuring presence in virtual environments: A presence questionnaire," *Presence Teleoperators Virtual Environ.*, vol. 7, no. 3, pp. 225–240, 1998, doi: 10.1162/105474698565686.
[36] J. Brooke, "SUS-A quick and dirty usability scale," in *Usability evaluation in industry*, P. W. Jordan, B. Thomas, B. A. Weerdmeester, and A. L. McClelland, Eds. London: Taylor & Francis, 1996.
[37] M. Slater, M. Usoh, and A. Steed, "Depth of Presence in Virtual Environments," *Presence*

Teleoperators Virtual Environ., vol. 3, no. 2, pp. 130–144, Jan. 1994, doi: 10.1162/pres.1994.3.2.130.

[38] P. E. Wais, M. Arioli, R. Anguera-Singla, and A. Gazzaley, "Virtual reality video game improves high-fidelity memory in older adults," *Sci. Rep.*, vol. 11, no. 1, pp. 1–15, 2021, doi: 10.1038/s41598-021-82109-3.

[39] Y. S. Shin, R. Masís-Obando, N. Keshavarzian, R. Dáve, and K. A. Norman, "Context-dependent memory effects in two immersive virtual reality environments: On Mars and underwater," *Psychon. Bull. Rev.*, 2020, doi: 10.3758/s13423-020-01835-3.