IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received October 26, 2018, accepted December 2, 2018, date of publication December 19, 2018,
date of current version January 7, 2019.

Digital Object Identifier 10.1109/ACCESS.2018.2886627

Augmented Reality Based on SLAM to Assess Spatial Short-Term Memory

FRANCISCO MUNOZ-MONTOYA[1], M.-CARMEN JUAN[1],
MAGDALENA MENDEZ-LOPEZ[2], AND CAMINO FIDALGO[2]
[1]Instituto Universitario de Automática e Informática Industrial, Universitat Politècnica de València, 46022 Valencia, Spain
[2]Departamento de Psicología y Sociología, IIS Aragón, Universidad de Zaragoza, 44003 Teruel, Spain

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

This work was funded mainly by the Spanish Ministry of Economy, Industry and Competitiveness through the AR3Senses project (TIN2017-87044-R) and cofinanced by the ERDF (European Regional Development Fund). Other financial support was received from the Gobierno de Aragón (Departamento de Innovación, Investigación y Universidad) and FEDER ''Construyendo Europa desde Aragón'' for the research group S31_17D.

ABSTRACT Spatial short-term memory is defined as the limited ability of people to retain and remember the location of elements for short periods of time. In this paper, we present the first AR app based on SLAM (Simultaneous Localization and Mapping) to assess spatial short-term memory. A total of 55 participants were involved in a study for remembering the real place where four virtual objects were located in the real environment. The participants were divided into two groups: the ARGroup (the participants learned the location of the virtual objects in the real environment in an adaptation phase using AR) and the NoARGroup (the participants learned the location of the objects by looking at photographs). The results indicated that the performance outcomes in remembering objects and their location for the participants in the ARGroup were statistically significantly greater than those obtained by the participants in the NoARGroup. From this result and our observations, we can conclude that touring the augmented environment helped the participants to better remember the location of virtual objects added to the real scene compared to looking at photographs of the environment. Furthermore, statistically significant differences were not found in relation to gender or age. Finally, our app has several advantages: 1) Our app works in any environment and does not require adding real elements to the environment; 2) the evaluators can select any real environment and place the virtual elements where they want and even change them between sessions; and 3) our app could work similar to the way spatial memory does in everyday life.

INDEX TERMS Augmented reality, cognition, mobile applications, psychology, SLAM, spatial memory.

I. INTRODUCTION

In the last two years, hardware (e.g., Microsoft HoloLens and Magic Leap One$^{\text{TM}}$) and software (e.g., Tango SDK, ARCore, and ARKit) have been developed with enough power to create Augmented Reality (AR) apps based on SLAM (Simultaneous Localization and Mapping) for mobile platforms. SLAM is a markerless tracking technology that tracks the environment without the need for adding any physical objects to the environment (e.g., markers or image targets). The SLAM mapping process obtains spatial data (e.g., 3D point clouds) of the environment to build a global reference map while simultaneously tracking the position of the subject [1]. AR based on SLAM offers many possibilities, and indoor location is undoubtedly one of them.

AR has not been exploited for processes of spatial orientation and location. The spatial ability of a person refers to the ability to solve spatial problems, such as perceiving distances and directional relationships, mentally transforming objects with respect to their position in space, locating elements in space, etc. It is important to know how the person is spatially oriented and what the factors are that influence this ability for the implications in daily and working life.

AR apps can be used for both assessment and training of spatial memory. As assessment tools, they enable the identification of alterations in spatial memory in both children and adults, e.g., determining if a person has difficulties that may affect his/her independence [2]. As training tools, they can improve human performance in situations involving spatial orientation, e.g., practicing aid strategies [3]. Improving spatial capabilities not only benefits orientation behavior, but it also has a positive impact on the recovery of other areas, such as motor and social relationships [4]. From a psychological point of view, knowing which variables are related to the performance obtained in AR apps will allow us to develop improved future designs oriented to spatial assessment and training with AR apps.

In this paper, we present a SLAM-based AR app to assess spatial memory. The objective of our work was to develop and validate our AR app for the assessment of spatial short-term memory by comparing the participants' outcomes in two different conditions: the ARGroup (learning the location of objects by using AR) vs. the NoARGroup (learning the location of objects by observing photographs of the environment). To our knowledge, there is only a single task that has been tested in two studies that has used AR (based on image targets, not SLAM) on mobile devices [5], [6]. Juan *et al.* [5] and Mendez-Lopez *et al.* [6] demonstrated that AR allows the development of applications that evaluate spatial ability while the person is moving. Using AR in any real environment (not limited to an area controlled by elements added to the scene) has great potential in the study of spatial orientation. This would allow the tasks to be more similar to those of everyday life. An important difference with these works [5], [6] is that our app can be used in any indoor environment and it does not require the inclusion of additional elements to the real scene. Moreover, the evaluator can personalize the environment by placing the virtual objects in the desired place. The primary hypothesis of our work was that the performance outcomes for remembering the objects in their location would be significantly greater for the ARGroup.

The paper is organized as follows. Section II focuses on the state of indoor positioning and the assessment of spatial memory assisted by computer. Section III presents the design and development of our app. It also briefly explains the hardware and software used to develop and run the app. Section IV details the characteristics of the participants involved in the study, the measures and the configuration of the environment used, and the protocol followed. Section V presents the results. Section VI discusses our work and results. Section VII presents our conclusions and identifies areas for future research.

II. BACKGROUND

A. INDOOR POSITIONING

The technologies that are currently available for indoor positioning (Indoor Positioning System (IPS)) are the following (https://www.infsoft.com/solutions/basics/whitepaper): WiFi, Bluetooth, VLC (Visible Light Communication), and UWB (Ultra Wide Band). In exceptional cases, the use of GPS is possible, but this technology does not work when there is no visual contact with several GPS satellites. GPS accuracy can vary between 5 and 20 meters. In many cases, since there are already WiFi access points in many buildings, WiFi positioning systems (WPS) can be installed. The method known as fingerprinting is used for positioning. The accuracy of this method varies between 5 and 15 meters. Another possibility is the use of Bluetooth beacons (BLE Beacons). With this method, the accuracy varies between 1 and 3 meters. The use of WiFi and Bluetooth are two technologies that have proven to be useful for indoor location. One technology still to be exploited is VLC. In this case, special LEDs and infrared lamps emit imperceptible flickering light, which is detected by the camera of a mobile phone or other sensor. In this case, the accuracy can be less than 50 cm. Recently, the company Decawave (https://www.decawave.com) introduced a technology based on ultra wide band radio (UWB) signals that can reach an accuracy of 10 cm., both indoors and outdoors. By placing a series of beacons, the position of a node can be located with high precision. This technology is based on the ability to measure the propagation time of the radio signal (and, therefore, the distance) between the elements of the system. According to the manufacturer's specifications, the range can be up to 40 meters through walls, and 300 meters in direct vision. However, our experience is that the presence of metal objects, water, and even people affects accuracy.

If the location is required to augment the scene (AR), none of the above-mentioned technologies offer the accuracy required to achieve an acceptable static error in the registration of augmented objects (placing the virtual object in the real scene with accuracy). The alternative is to use the SLAM technique for indoor positioning. The use of SLAM with mobile devices offers many possibilities. One of the most important ones is that it is wireless. The users have freedom of movement since movement is not limited by cables. Moreover, it is not necessary to add other elements to the environment such as beacons.

Recently, much attention has been paid to the possibilities of SLAM for indoor positioning. For example, Rehman and Cao [7] used the Metaio SDK (a framework for marker and SLAM tracking) for indoor tracking. Using the Metaio SDK, they scanned the environment that consisted of visual features (3D point clouds) that were stored as trackables. Those trackables were associated with their corresponding locations and navigation-related information. The camera and inertial sensors of the device were used to track the 3D point clouds and device orientation. They conducted a study for indoor navigation to compare the performance of the participants using a HMD (Google Glasses), a Smartphone (Samsung Galaxy S4), and a paper map. They found that both digital navigation tools were better than the paper map in terms of shorter time and lower workload, but the digital aids resulted in worse route retention. Polvi *et al.* [8] presented SlidAR, which is a 3D positioning method for SLAM-based handheld AR. SlidAR utilizes 3D ray-casting and epipolar geometry for virtual object positioning. They conducted a study that involved 23 participants. They compared the efficiency of the SlidAR method against a device-centric positioning method. Their results showed that SlidAR was significantly faster, required significantly less device movement, and got significantly better subjective evaluations from the participants. SlidAR offered somewhat higher positioning accuracy. Piao and Kim [9] presented an adaptive monocular visual–inertial SLAM method for real-time AR applications in mobile devices. Their results demonstrated the effectiveness of performance improvement using their proposed method (up to 18.8%). Egodamage and Tuceryan [10] presented a collaborative AR framework based on distributed monocular visual SLAM.

Besides the possibilities for indoor positioning, SLAM-based AR can be used in many other applications. For example, Chen *et al.* [11] presented an efficient and effective 3D surface reconstruction framework for an intra-operative monocular laparoscopic scene. They checked the accuracy of the camera tracking by comparing the results of the video camera tracking with the recorded ground-truth camera trajectories. Root mean square errors of 1.24 mm and 2.54 mm. were obtained for the camera trajectories and the surface reconstruction, respectively. Their results show the potential of AR based on SLAM to be used in minimally invasive surgery.

B. ASSESSMENT OF SPATIAL MEMORY ASSISTED BY COMPUTER

Spatial short-term memory can be defined as the limited ability of people to retain and remember the location of elements for short periods of time [12]. Virtual Reality (VR) or AR applications allow objective indicators of a person's spatial learning to be obtained through a presentation of stimuli (varied and diverse) and the storage of responses (reaction times, successes-failures, distance traveled, speed, etc.) (e.g., Picucci *et al.* [13]; Juan *et al.* [5]; Walkowiak *et al.* [14]). Applications of this type suppose an advantage with respect to the evaluation and training of the person in a natural environment (temporal and economic costs, etc.). Therefore, the use of VR in the study of human spatial ability is becoming more frequent. Traditionally, VR applications were used in procedures that were developed in simple natural environments (rooms, laboratories, etc.). In these scenarios, the person interacts in a more or less complex virtual environment without physical displacement such as rooms where the person is sitting in front of a computer screen and performs a task exploring a VR environment (e.g., Picucci *et al.* [13]; Cimadevilla *et al.* [15]; Walkowiak *et al.* [14]). However, the physical displacement component is important in spatial ability [16]. Following this idea, the latest VR works have incorporated physical displacement (Rodríguez-Andrés *et al.* [17], [18]; Cárdenas *et al.* [19]).

Rodríguez-Andrés *et al.* [17] presented a VR task for the assessment of spatial short-term memory. In this work, they examined the influence of the type of interaction used on the ability to recall the place of the objects and the perceived usability and satisfaction of the children with the task. They used a large screen (120") for the visualization. Two interaction modes (the physical active condition vs. the physical inactive condition) were examined. For the physical active condition, they used a Wii mote and a Wii balance board. For the physical inactive condition, they used a gamepad. A total of 160 children participated in their study. There were no statistically significant differences in the results of the task using the two types of interaction. They found correlations between the scores obtained using their VR task and a traditional procedure for assessing spatial short-term memory. Their results revealed that the type of interaction used did not affect the performance of children in the VR task.

Cárdenas *et al.* [19] presented a VR task based on a maze that assesses spatial short-term memory in adults involving physical movement and immersion. As in previous works [17], [18], they used two different interaction types (the physical active condition vs. the physical inactive condition). For the physical active condition, they used a real bicycle. For the physical inactive condition, they used a gamepad. For immersion, they used a VR HMD (Oculus Rift). A total of 89 adults participated in their study. Their results showed that the performance on their task was better in the participants who used the physical inactive condition. Usability and satisfaction did not differ between conditions. The performance on the task correlated with the performance on other classical neuropsychological tests for the assessment of short-term memory and spatial memory.

With regard to physical displacement, AR offers new possibilities. To our knowledge, only a single task tested in two studies has used AR for the assessment of spatial orientation [5], [6]. Their task assessed the ability of the participants to remember the location in the real world of an increasing number of virtual objects that appeared augmented in the real world. The AR app used image targets. The testing area was a square of about five meters on each side. The testing area was surrounded by light brown paper to a height of 1.5 meters. The boxes were distributed in a circle with a radius of 1.85 meters. The image targets were placed inside real boxes, which served as locations. These boxes were strategically located in the testing area. The task consisted of seven different levels. The number of boxes ranged from 2 to 14, depending on the level. Their study involved 76 children divided into two groups: preschool (5–6 years old) and primary school (7–8 years old). They obtained significant performance outcomes in the AR task in favor of the older group. They found significant correlations between traditional tests and scores for the AR task. Their study revealed that the younger children were more satisfied with the AR task. As mentioned above, our work goes one step further in demonstrating the potential of SLAM-based AR to assess spatial memory.

III. DESIGN AND DEVELOPMENT

In this section, we describe the phases of the app. We also detail the hardware and software used.

A. THE APP

The central part of the app is a task that allows the participants to tour a real environment in which they must look for virtual objects and remember their location. For this purpose, the person in charge of the evaluation must first (and only once) configure the scene for the task in two phases:

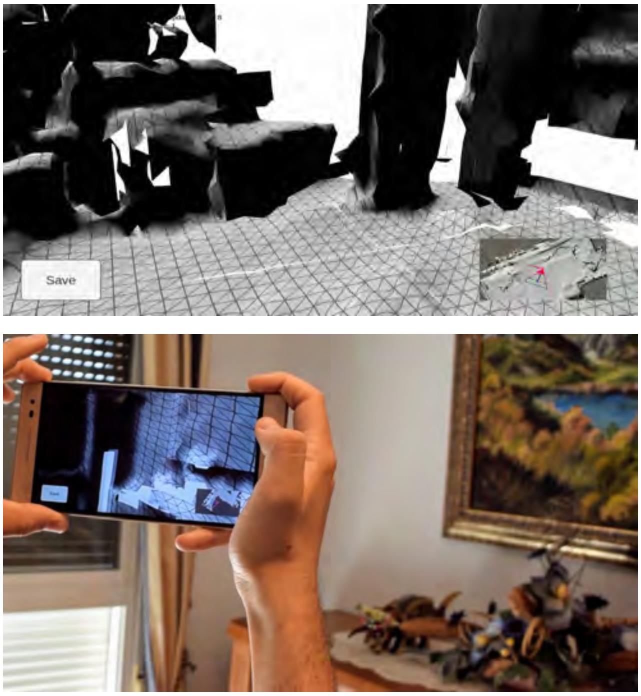

FIGURE 1. **Scanning phase. View of the environment used in the study.**

1) environment scanning; and 2) object configuration. The phases for configuring the scene for the task are the following:

1) CONFIGURATION PHASE. THE ENVIRONMENT SCANNING PHASE

This is a configuration phase in which the supervisor scans the real environment where the validation of the task will be performed (Fig. 1). This scan is necessary in order to store the morphology of the environment in the device.

2) CONFIGURATION PHASE. THE OBJECT CONFIGURATION PHASE

Using the information stored in the previous phase, the supervisor locates the different 3D objects in the real scanned environment. A total of four objects are available for positioning. These four objects are: a sculpture, a telephone, a fountain pen, and a toy car. The supervisor can place these four objects in the desired locations using the strategy she/he chooses. The objects should be placed on plane surfaces (horizontal or vertical). The orientation of an object is obtained by the normal vector of the identified surface and the direction from the camera to the object. The forward face of the object is facing the position of the supervisor, who handles the device at each moment. The up vector of the object has the same direction as the normal vector of the detected surface. The supervisor has the freedom to choose the spot on a surface where the object will be placed (with two degrees of freedom (x and y)). For example, if the detected plane is on a horizontal table, the supervisor has the possibility of moving the object forward and backward (Y axis) as well as to the right and to the left (X axis) on the table. The supervisor will not be allowed to move the object up and down (Z axis). The bottom side of the object is attached to the detected surface. The supervisor will not be allowed to rotate the object. The objects are always facing the supervisor's position, and their up vectors are perpendicular to the plane of the table. During this phase, the supervisor takes photographs of the environment where the virtual objects are located. These photographs will be used in the memorization phase.

The phases of the task for the assessment of spatial memory that should be performed by the user are the following (see the Video File for a video demonstrating the functionality of our app):

1) USER'S PHASE. THE ADAPTATION PHASE

This phase has two objectives. The first objective is to provide an initial experience with the mobile device and the task so that the user becomes familiar with them. The user learns how to hold the device, how to move inside the virtual environment, and how it works. Two conditions are used:

A. The AR Condition in which the virtual elements appear overlapped in the real environment (AR). The users can approach the virtual objects as close as they like and view them from different angles while familiarizing themselves with the environment, the device, and the task.
B. The NoAR Condition in which the users only see the real environment through the device screen, but without virtual elements.

These two conditions were defined in order to corroborate that the performance outcomes for remembering the virtual objects in their location would be significantly greater for the AR Condition (the primary hypothesis). In other words, our hypothesis is that seeing the virtual objects integrated in the real environment helps to memorize them. This memorization would be statistically greater when compared to not seeing the virtual objects in the real environment and memorizing their location using only photographs.

During this phase, the users are asked to inspect the environment looking for virtual objects that are not visible in the real environment. These objects are the objects that were added in the object configuration phase. These virtual objects appear in the ARCondition (Fig. 2), but they do not appear in the NoAR Condition (Fig. 3). As can be observed in Figs. 2-3, a virtual red car appears on the table in the AR Condition (Fig. 2), but it does not appear in this location in the NoAR Condition (Fig. 3). The users have a total of two minutes to complete this phase. They do not receive any help regarding where the objects are, the objects that they have seen, or those that they must see.

2) USER'S PHASE. THE MEMORIZATION PHASE

During this phase, the photographs taken in the configuration phase are shown (Figs. 4-5). These photographs show the virtual objects in the real environment. During this phase,

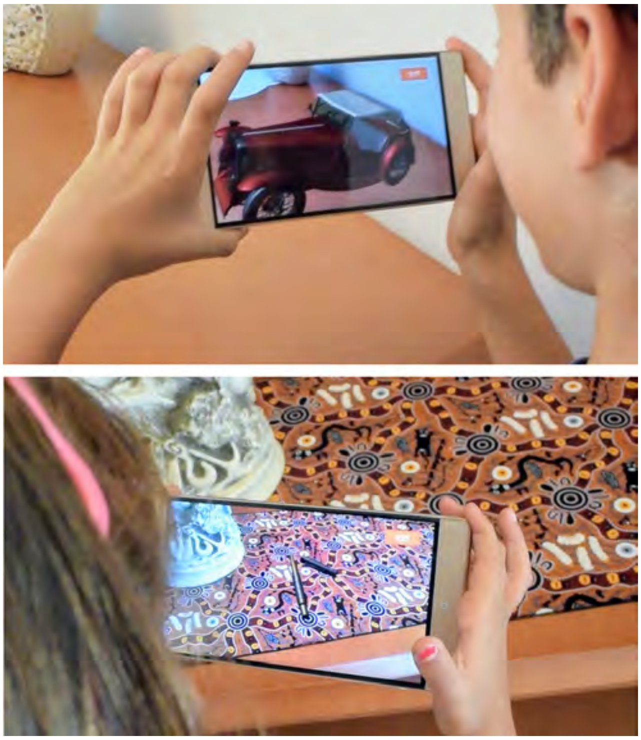

FIGURE 2. Two examples of children performing the adaptation phase in the AR Condition.

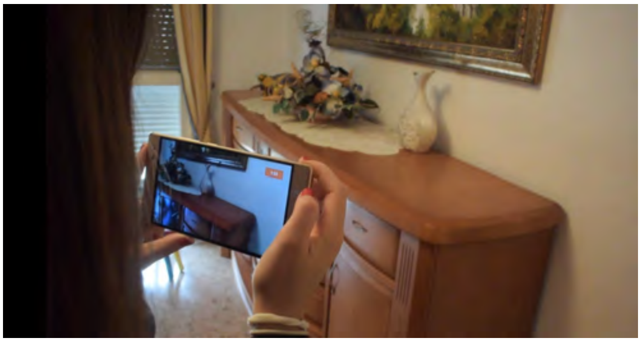

FIGURE 3. An example of a participant performing the adaptation phase in the NoAR Condition.

the users must memorize the location of the different objects in the real space. This phase is the same for the two groups (AR and NoAR). However, depending on the condition used in the adaptation phase, the users may have already memorized these locations (the AR Condition) or it may be the first time that they see the objects through 2D photographs (the NoAR Condition). In other words, the participants in the AR Condition may have already seen the objects mixed with the real world in the adaptation phase. This memorization phase could serve to reaffirm the information perceived in the adaptation phase. However, the participants in the NoAR Condition had not seen these objects because they did not appear in the real world when they were touring with the mobile device in the adaptation phase. Since this is the first time that these participants see the virtual objects, they must memorize these virtual objects and relate them to their real-world locations. The users can zoom in or zoom out of the photographs as many times as they wish. The participants have a total of one and a half minutes to complete this phase. The participants in the AR Condition did not pay as much attention as the participants in the NoAR Condition. However, in this study the time for the memorization phase was fixed and the participants could not leave before the fixed time.

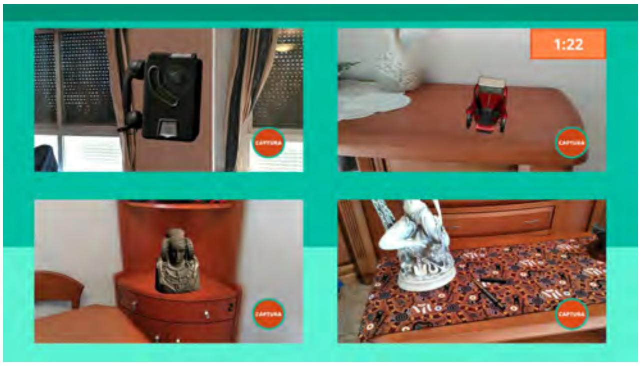

FIGURE 4. Memorization phase. The photographs taken in the configuration phase are shown on the screen device.

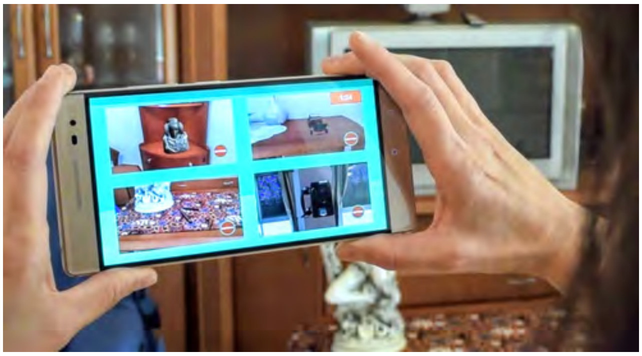

FIGURE 5. Memorization phase. A user performing the memorization phase.

3) USER'S PHASE. THE EVALUATION PHASE

This phase assesses the users' ability to remember the location of the virtual objects in the real environment learned in the previous phases. The users are asked to locate the virtual objects in their correct locations. The users can select the object to be located among the four virtual objects that appear in a selection bar with buttons on the right side of the screen (Fig. 6). When the button of a given object is selected, it is automatically positioned in the center of the device screen. The object adapts to horizontal and vertical surfaces. In other words, the app identifies the horizontal and vertical surfaces and allows objects to be placed on them. The functionality for the placement of the objects by the participants is the same as that described in the object configuration phase. When the object is on the desired surface, the user presses the ''Set'' button and the object is anchored in its current position (Fig. 7). The objects do not have to be placed in exactly the

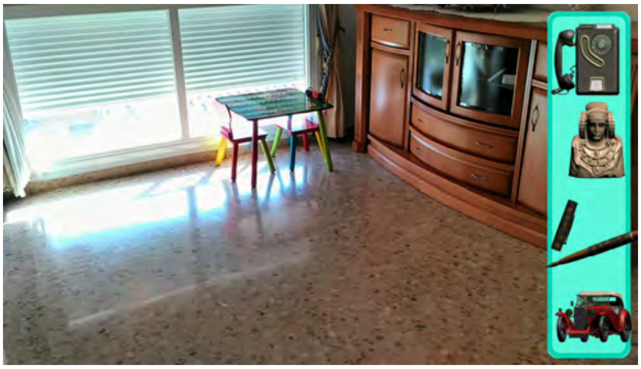

FIGURE 6. Evaluation phase. The selection bar with 4 buttons showing the images of the objects to be placed in the real environment on the right side of the screen.

FIGURE 7. Evaluation phase. The car is placed on the correct table.

same location. There is a margin of error of a sphere of half a radius meter from the point where the object should be placed. The users are informed whether or not they have positioned the object correctly. The users have three attempts to position each object correctly. On the last attempt, the object remains fixed in the position where it was placed, informing the user whether it was a success or failure, and then disappears from the selection bar. The users do not get any help regarding where to place the objects.

B. DEVELOPEMENT

1) HARDWARE & SOFTWARE

The device used for the development of the app and the study was a Tango smartphone, Lenovo Phab 2 Pro. The main characteristics of the mobile phone are: Dimensions (179.8 × 88.6 × 10.7 mm); Weight (259 grams); Size (6.4 inches); and Resolution (1440 × 2560 pixels). This is one of the two devices available on the market that can run apps developed with Tango SDK. The main feature that provides this device with such special functionality is that it has three built-in cameras (a color camera with 16 MP, a wide-angle camera, and a depth camera). These three cameras and the Tango SDK provide the device with SLAM technology. The three main functionalities are: motion tracking, learning area, and depth perception.

The app was developed using the Unity3D game engine (https://unity3d.com). For our app, Unity3D offers two advantages: the abstraction of the life cycle of a graphic application, and the integration of the Tango SDK.

2) USER INTERFACE

An intuitive interface was designed with a minimal number of buttons and elements in order not to distract the user from the task. The interface changes during the process of the task, while still maintaining a common thread. During the adaptation and evaluation phases, the users hold the mobile device in front of their body while the image captured by the camera is displayed on the screen. In the adaptation phase with the AR Condition, the virtual elements overlap this image. Additional buttons are incorporated in the evaluation phase. As already mentioned, a selection bar with different buttons appears on the right side of the screen with the images of the objects to be placed in the real environment (Fig. 6). The process of positioning the object in the real environment begins once one of these buttons is selected.

Different options were considered to facilitate a stable positioning by the user. One proven option was to drag objects across the screen and drop them into place using touch. However, this method becomes difficult when the user has to hold the device with one hand and drag the objects with the other. Since precise positioning was very difficult to achieve with this method, we chose to place the object in the middle of the screen and adapt its position to the surface at the corresponding point. The object adapts its position to the surfaces in the center of the screen as the user moves around with the device. A "Set" button was incorporated. This button is pressed by the user when the object is in the desired position and the object is anchored in its place.

In the memorization phase, the interface consists of a frame showing the four different photographs of the objects positioned in the real environment. The user is able to zoom in and out by pressing on them.

3) AR SERVICE

For the development of the app, we used three features offered by the Tango SDK:

- **Area Learning**. This feature allows the environment to be scanned. It performs an initial scan of the environment in which the scene morphology is stored in the device. The SLAM technique is used to find and store visual features that enable a future location of the device in the environment. With the scanned and saved environment, it is possible to indicate certain points where the virtual elements will be located. These points in the space are stored along with the characteristics of the environment.
- **Motion Tracking**. The device knows its relative position with respect to the real environment at any moment. This feature allows the mobile device to show the virtual objects in their proper place and even trace the path that

TABLE 1. Gender and age distribution of the participants.

	CHILDHOOD (<=15)	YOUTH (16-25)	*ADULTHOOD (26-50)*	MIDDLE AGE (>50)
Men	10	4	13	4
Women	0	3	17	4

users follow. This feature is used in the adaptation and evaluation phases.

- **Depth Perception**. The depth camera is used to detect flat surfaces in the space on which to place virtual objects. To obtain these flat surfaces, the app generates a cloud of points of the environment that appears on the screen. This cloud of points is analyzed to find flat surfaces. The flat surfaces that we use can be horizontal and vertical.

4) STORAGE OF DATA

During the adaptation and evaluation phases, the app stores different data. We used some of this data in our analysis. This data includes: the user ID, the condition of the adaptation phase (AR or NoAR), the time spent on each phase, the successes achieved in the evaluation phase, and the relative position of the user with respect to the environment. The data is collected without altering the execution of the task.

IV. DESCRIPTION OF THE STUDY

This section presents the characteristics of the participants involved in the study, the measurements used, the configuration of the environment, and the steps followed.

A. PARTICIPANTS

A total of 55 subjects, ranging in age from 8 to 72 years old, were involved in the study. The mean age was 36.53 ± 15.78 years old. There were 31 men (56.36%) and 24 women (43.63%). Table 1 shows the participants' distribution for gender and age. The participants or their parents were informed about our study and their objectives. They signed a written consent form. The principles expressed in the Declaration of Helsinki were followed for all of the clinical research. The study and the written consent form were approved by the Ethics Committee of the Universitat Politècnica de València, Spain.

B. MEASUREMENTS

The app stored the following variables: the type of augmentation (AR or NoAR in the habituation phase), the errors committed in the evaluation phase, and the duration of the phases.

Presence in virtual environments can be defined as an individual and context-dependent user's response that is related to the experience of "being there" [20]. Witmer and Singer [21] define presence as a psychological state of "being there" mediated by an environment that engages our senses, captures our attention, and fosters our active involvement. However, according to Regenbrencht and Schubert [22], this definition cannot be applied exactly to AR. However, in AR, presence can also be achieved by measuring the presence of virtual elements in the real environment [22]. To measure the sense of presence, we added ten questions that are adapted from the Witmer and Singer questionnaire [21]. We also added two questions from the questionnaire proposed by Slater *et al.* [23]. These two questionnaires have commonly been used to measure presence in VR environments. For AR environments, we adapted five questions from the Regenbrech and Schubert questionnaire [22]. We included ten questions from the Witmer questionnaire (vs. 3.0, 4-factor model) [24] to measure presence. These 10 questions considered the four factors identified for their presence questionnaire. The numbers in parentheses are the number of the questions in the Witmer questionnaire version 3.0. These factors are: Involvement (2, 6, 18); Visual fidelity (15, 16); Adaptation/Immersion (20, 21, 24), and Interface Quality (19, 23). In total, we have seventeen questions to measure presence.

For perceived usability, we included six questions adapted from the SUS questionnaire proposed by Brooke [25].

To assess interest/enjoyment, perceived competence, and pressure felt, we included eleven question from the Intrinsic Motivation Inventory (IMI) [26]. Specifically, we included five questions for interest/enjoyment, five questions for perceived competence, and one question to measure the pressure felt. We also included two questions to measure the perceived mental effort and the physical effort in arms and hands.

For perceived satisfaction, we included four questions based on our previous experiences (e.g., [27]).

The questionnaire (39 questions) was filled out online in a web-based format. All of the questions were formulated in a positive manner. All of the questions used a 7-point Likert scale ranging from 1 "Totally disagree" to 7 "Totally agree".

C. CONFIGURATION OF THE ENVIRONMENT

The study was carried out in a room of 42 square meters. The room had the furniture commonly found in a dining room. The virtual objects could be mimicked in that environment. The four virtual objects were positioned in the room. Fig. 8 shows the shape of the room and the location of the four virtual objects.

D. PROCEDURE

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

- The ARGroup: Participants who learn the location of the virtual elements that are overlapped in the real environment in the adaptation phase using AR.
- The NoARGroup: Participants who see the real environment through the device screen, but without virtual elements. These participants learned the location of the objects by looking at photographs in the memorization phase.

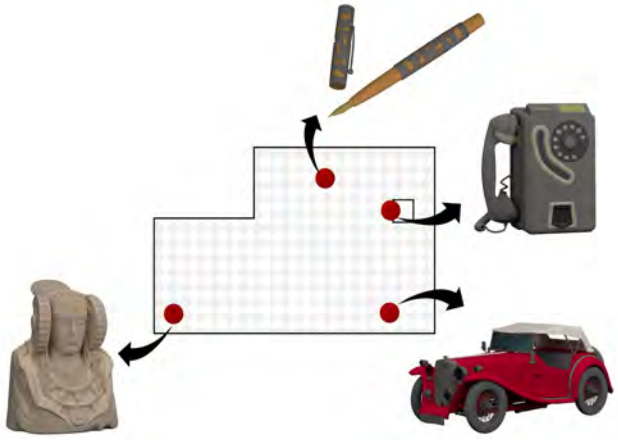

FIGURE 8. **Study environment: shape of the room and location of the four virtual objects.**

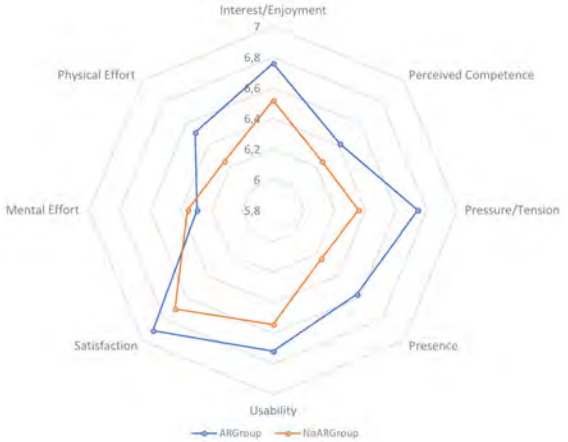

FIGURE 9. **Graph showing the mean scores for the different variables measured in the questionnaire.**

The protocol was the following:

1) The participants performed the task (ARGroup or NoARGroup).
2) The participants filled out a questionnaire.

V. RESULTS

This section details the analysis carried out with the data collected during our study. An initial descriptive analysis was carried out to explore means, standard deviations, and other measurements. Data normality was checked, and the appropriate statistical tests were applied. To check data normality, we applied the following tests: Shapiro-Wilk ($W = 0.591$, $p < 0.001^{**}$), Kolmogorov-Smirnov ($D = 0.428$, $p = 0.001^{**}$), and Anderson-Darling ($A = 11.305$, $p < 0.001^{**}$). The three tests indicated that our sample did not fit the normal distribution. Therefore, non-parametric tests were used (the Mann-Whitney U test and the Spearman correlation for correlation tests). All of the tests are presented in the format (statistic U/W, normal approximation Z, p-value, r effect size). The symbol ** indicates the statistical significance at level $\alpha = 0.05$. The statistical open source toolkit R (http://www.r-project.org) was used to analyze the data.

A. PERFORMANCE OUTCOMES

In order to know how the use of AR for learning the location of different objects affects the performance outcomes of the participants, we compared the performance outcomes between the two groups (the ARGroup vs. the NoARGroup). The score variable was created by counting the number of objects placed correctly for the four objects.

To determine whether or not there were differences in remembering and placing objects in their correct location between the participants of the two groups (AR (3.926 ± 0.267) vs NoAR (3.25 ± 0.799)), a Mann-Whitney U test was applied ($U = 187$, $Z = 0.428$, $p < 0.001^{**}$, $r = 1.143$). This result showed that there was a statistically significant difference between the two groups in favor of the ARGroup (the group that learned using AR). From this result, we can deduce that AR helped users to better learn the position of objects in the environment.

B. SUBJECTIVE PERCEPTION AND SATISFACTION OUTCOMES

The online questionnaire described in Section IV.B (Measurements) was used to measure the subjective perception of the participants with the task and their performance. The questions in that questionnaire were grouped in variables to measure different factors. The means of these variables are shown in Fig. 9. This figure shows that all of the mean scores were quite high. The lowest value was 6.23 in presence for the NoARGroup. Our explanation for this result is that the participants did not visualize the virtual objects in the adaptation phase.

A Mann-Whitney U test was applied to the defined variables and the two groups (AR vs. NoAR). There was a statistically significant difference in the satisfaction experienced by the users between the two groups in favor of the ARGroup ($U = 255$, $Z = 0.356$, $p = 0.0148^{**}$, $r = 0.57$). There was a statistically significant difference in presence experienced by the users between the two groups in favor of the ARGroup ($U = 247$, $Z = 0.189$, $p = 0.0271^{**}$, $r = 0.599$). There were no statistically significant differences for the other variables. We would like to add that the means for all of the variables and for the two groups were above 6 on a scale from 1 to 7. The means for all of the variables except one (Mental Effort) in the ARGroup were higher than in the NoARGroup (e.g., the mean for Satisfaction in the ARGroup was 6.907). These means and the analysis performed demonstrate the positive perception of our task by the participants.

Since statistically significant differences were observed in the satisfaction and presence variables with respect to both groups, correlations between variables were applied separately. A Spearman correlation was applied to determine if there is a significant correlation between some of the measured variables in each of the two groups (the ARGroup

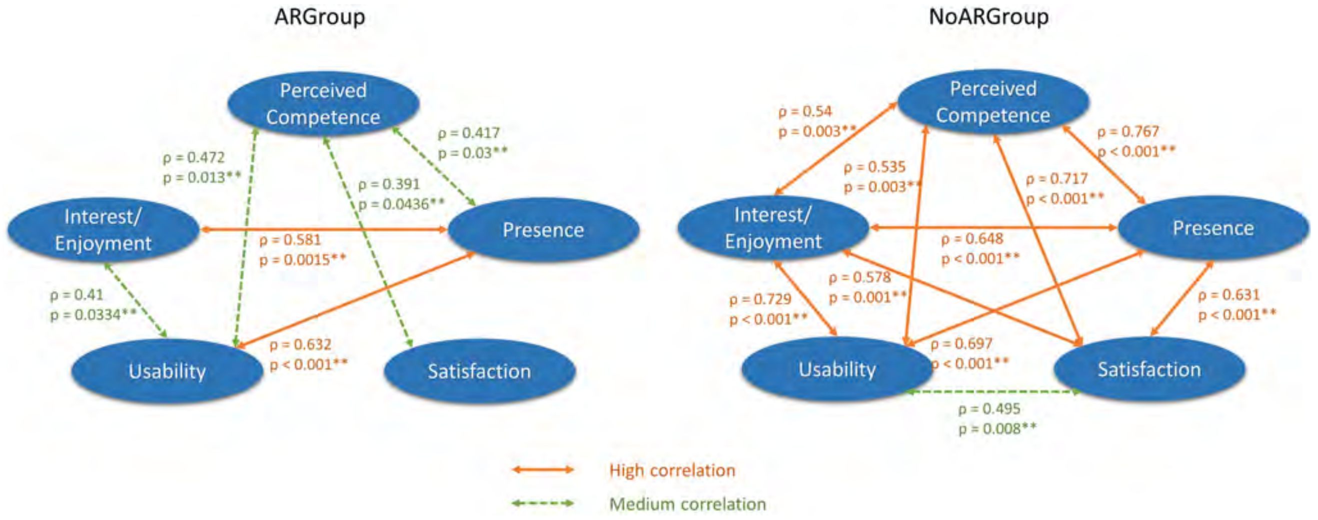

FIGURE 10. The correlation plot among the analyzed variables for the two groups (the ARGroup and the NoARGroup).

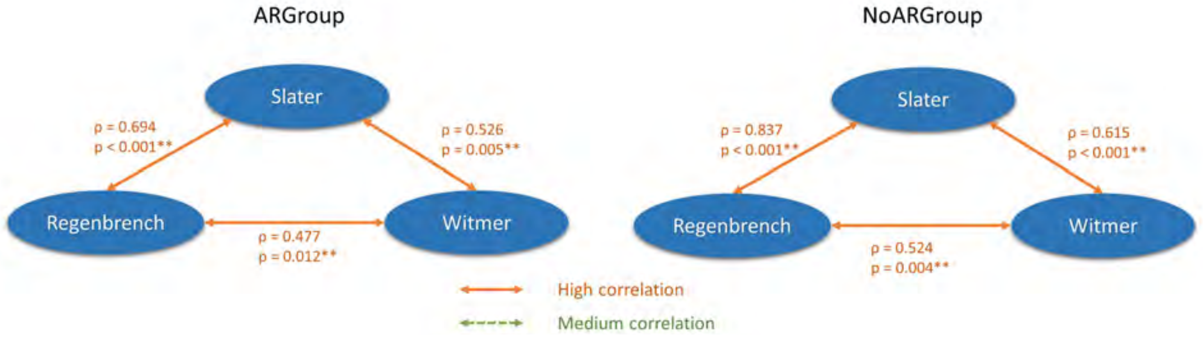

FIGURE 11. The correlation plot for the variables related to the presence questionnaires of Slater, Regenbrench, and Witmer for the two groups (the ARGroup and the NoARGroup).

vs. the NoARGroup). Fig. 10 shows the correlation plot. There are six significant positive correlations in the ARGroup and there are ten significant positive correlations in the NoARGroup.

In order to measure presence, questions based on the questionnaires of Slater, Regenbrench, and Witmer were used. A Spearman correlation test was applied to the variables related to these three questionnaires for the two groups (Fig. 11). Fig. 11 shows that there are significant positive correlations among the variables related to the three questionnaires in the two groups. This is a good result because it indicates that the selected and adapted questions of these three questionnaires were appropriate for measuring the level of presence in our task.

The different factors identified in the Witmer questionnaire were also analyzed. Fig. 12 shows the correlation plots for the two groups. In the work of Witmer *et al.* [24], four significant positive correlations were obtained (Involment≪-≫Adaptation/Immersion; Adaptation/Immersion≪-≫Sensor Fidelity; Sensor Fidelity≪-≫ Involvement; Adaptation/Immersion≪-≫Interface Quality). In our study, we found five significant correlations in the ARGroup and the NoARGroup. Three of them were the same as in the Witmer study (Involment≪-≫Adaptation/Immersion; Visual Fidelity≪-≫Involvement; Adaptation/Immersion≪≫Interface Quality). The other two significant correlations were: Involment≪-≫Interface Quality and Visual Fidelity≪-≫Interface Quality.

C. GENDER AND AGE COMPARISONS

A Mann Whitney U test was applied to check if gender affected the score. The results indicated that there was no statistically significant difference in gender (U = 385.5, Z = 0.278, p = 0.792, r = 0.037). To determine if the score obtained by the participants was affected by age, we applied a Kruskal Wallis test. The results showed no statistically significant differences for the age factor ($\chi^2(3)$ = 0.986, p = 0.805, r = 0.107). For the variables of the questionnaire (Perceived Competence, Interest, Presence, Pressure, and Usability), Mann Whitney U tests were applied and no statistically significant differences were obtained regarding gender.

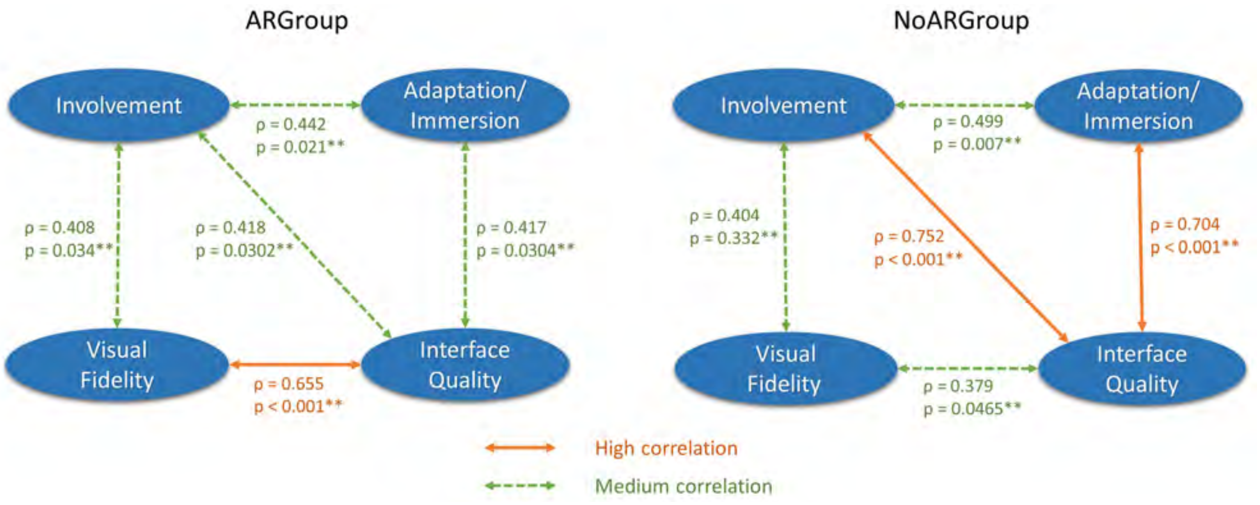

FIGURE 12. The correlation plot for the factors identified by Witmer for the two groups (the ARGroup and the NoARGroup).

With regard to the participants' experience and their age, we checked if there were differences in the experience that the users had during their performance due to their age. A Kruskal Wallis test was applied to each variable. The results showed that there were two variables that offered statistically significant differences depending on the age of the participants. These variables were: Interest ($\chi^2(3) = 9.003$, $p = 0.029^{**}$, $r = 0.303$) and Usability ($\chi^2(3) = 19.298$, $p < 0.001^{**}$, $r = 0.532$). For Interest, there was a statistically significant difference between children and the rest of the groups and no significant difference between groups who were over 15 years old. Mann-Whitney U tests were applied and the results were: Childhood vs. Youth ($U = 13$, $Z = 2.275$, $p = 0.0186^{**}$, $r = 0.552$); Childhood vs. Adulthood ($U = 66$, $Z = 2.746$, $p = 0.005^{**}$, $r = 0.434$); Childhood vs. Middle Age ($U = 16.5$, $Z = 2.152$, $p = 0.031^{**}$, $r = 0.507$). For the Usability variable, the results after applying the Mann-Whitney U tests were similar to the results obtained for the Interest variable. In other words, there were statistically significant differences between the group of children and the rest of the groups. Childhood vs. Youth ($U = 0$, $Z = 3.467$, $p < 0.001^{**}$, $r = 0.841$); Childhood vs. Adulthood ($U = 35.5$, $Z = 3.656$, $p < 0.001^{**}$, $r = 0.578$); Childhood vs. Middle Age ($U = 11.5$, $Z = 2.548$, $p = 0.00873^{**}$, $r = 0.601$). There were no statistically significant differences between the groups who were over 15 years old. Fig. 13 shows a box plot where the values for the Usability variable and different age groups are shown graphically. The box plot for the Interest variable shows a similar trend.

VI. DISCUSSION

In this paper, we have presented a SLAM-based AR app to support the assessment of spatial memory. The central part of the app is a task that allows participants to tour a real environment, in which they must to search for virtual objects and remember their location. To our knowledge, only one task has been tested in two studies that used AR for the assessment of spatial memory [5], [6]. However, those studies did not use SLAM-based AR. Those studies used AR based on fiducial markers (image targets) added to the real environment. Our work goes one step further in demonstrating the possibilities of SLAM-based AR for assessing spatial memory. In this paper, we carried out a study comparing the effects of using and not using AR for learning where the different objects were placed in the real environment. We also tried to determine whether using AR creates a significant difference in the user's experience. Our study involved 55 participants counterbalanced in AR vs. NoAR conditions (gender and age).

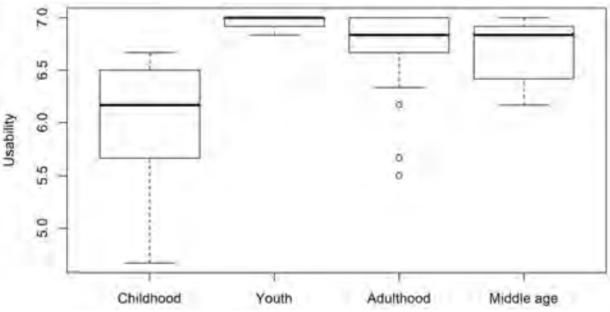

FIGURE 13. Box plot for the Usability variable and different ages.

The main difference between the two groups (the ARGroup and the NoARGroup) was that the participants of the ARGroup learned the location of the virtual objects placed in the real environment in the adaptation phase using AR. This phase allowed them to pay attention to details of the environment and thus facilitate more specific learning. These participants also observed the photographs of the environment with the virtual objects in the memorization phase. This phase was useful for participants to reinforce information about virtual objects and their location. Some of these participants were interested in this reinforcement, but other participants

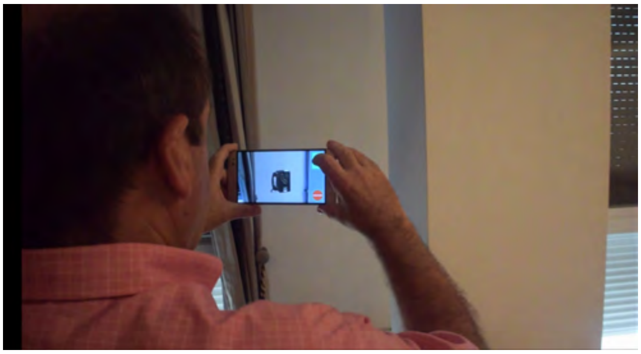

FIGURE 14. Participant of the NoARGroup placing the telephone on the wall and not on the pillar.

perceived it as being redundant and did not pay much attention. In contrast, the participants of the NoARGroup toured the real environment and could pay attention to details, but without seeing the virtual objects in the environment. These participants had to pay attention to the photographs of the environment with the virtual objects shown in the memorization phase. In this phase, these participants paid attention to the photographs and looked at the real environment in order to be sure of the position in which they should place the virtual objects.

The results of this study show that, for the participants in the ARGroup, there was a statistically significant difference in remembering objects and their location. This corroborates our main hypothesis. The difference was basically due to two objects, the fountain pen and the telephone. The fountain pen was placed on a table with a tablecloth that was the same as the one on another table in the environment. The participants of the NoARGroup confused the two tablecloths and some of them placed the fountain pen on the wrong tablecloth. This tablecloth is shown in Fig. 2. The telephone was placed on a pillar of the room that the users of the NoARGroup confused with the wall behind it. Fig. 14 shows a participant of the NoARGroup placing the telephone on the wall and not on the pillar. From these results and our observation, we can conclude that touring the augmented environment helped participants better remember the location of virtual objects that were added to the real scene. Moreover, if we take into account the successes when placing objects using the AR condition (3.926 ± 0.267), we can conclude that SLAM-based AR can be used for the development of apps to assess spatial orientation.

This result complements the results that we obtained in previous works [5], [6], which demonstrated that AR based on fiducials could also be used for the development of apps for the assessment of spatial memory.

When [5], [6] are compared with our work, our proposal has several advantages: 1) The app presented works in any environment and does not require adding real elements to the environment; 2) The evaluators can select any real environment and place the virtual elements where they want and even change them between sessions; 3) Our app could work in a way similar to the way spatial memory does in everyday life.

The app also has some advantages when compared to other methods of evaluating short-term memory: 1) The mental representation of the environment differs on the kind of space that is coded. Our app offers the possibility to investigate short-term spatial orientation by walking through the environment (i.e., the navigational space). The navigational space is the type of space in which many human behaviors take place. This is a great advantage compared to classical neuropsychological tests (e.g., Corsi Block Tapping Test [28]), which evaluate spatial memory in the physical environment within the reaching distance (i.e., near space). 2) VR systems have already been used to assess short-term spatial memory (e.g., Cánovas *et al.* [29]; Spieker *et al.* [30]). VR through Head-Mounted Displays tends to induce cybersickness [31]. None of the side effects attributed to cybersickness was experienced by any of the participants in our study.

The results also show that the performance outcomes were independent of the gender and age of the participants. This suggests that, regardless of gender or age, our task has proven to be suitable for assessing spatial memory. With regard to gender, our result is in line with the work in [5]. With regard to age, our result is different from the work in [5]. In [5], the study involved two groups of children, preschool (5–6 years old) and primary school (7–8 years old). Significant improvement outcomes were obtained with the task in the older group. The task had seven levels, with three trials in each level and with an incremental number of objects to remember in each level (from 1 to 7). In this regard, it is well-known that visuospatial short-term memory skills increase as the brain develops [32], [33]. Our argument for this result is that the level of complexity of the work in [5] is different from the work presented here. If the task were more complex, similar results could be obtained. Greater complexity could be incorporated and studied in future developments.

With regard to the experience with the app and the answers of the participants to the on-line questionnaire, the means rated for all of the variables except one (Mental Effort) in the ARGroup were greater than for the NoARGroup (Fig. 2). The participants in the ARGroup also experienced a statistically significant higher level of satisfaction and sense of presence. With regard to Mental Effort, our explanation for this result is that the participants in the NoARGroup did not have virtual objects to look for. Therefore, their mental effort was lower. The participants in the ARGroup had to tour the environment in order to search for virtual objects.

However, the differences in the users' experience regarding age but not gender were significant. No statistically significant differences were obtained for gender. Only two variables, Interest and Usability, showed statistically significant differences when children under 16 years old were compared with the rest of the groups. The scores for Interest and Usability were lower in the childhood group. These results are in contrast to previous works such as the work of Bacca *et al.* [34], which indicates that AR offers advantages such as ''motivation'', ''student engagement'', and ''improved perceived enjoyment''. Our explanation for the low score in Interest of the childhood group is that the children expected the task to be more entertaining and fun. They expected to play with a game similar to the ones that they are used to (e.g., soccer). If our task is to be used with this group, it should be customized so that the objects are more suitable for them and offer more playful activities. For Usability, all of the users handled the mobile device without any physical support. Sometimes, the users had to hold the mobile device with one hand and touch the screen with the other without covering up the cameras on the back of the device. This was more difficult for the children, especially the smaller ones. To facilitate the handling of the device and to provide more stability and safety, an external case could be designed and adapted [5], [35], [36]. The external case could be printed on a 3D printer [5]. This would also protect the device from damage.

For the subjective perception and satisfaction outcomes and their correlation plots (Fig. 10), there were six significant positive correlations in the ARGroup and there were ten significant positive correlations in the NoARGroup. Our explanation for these results is that the participants of the NoARGroup scored an average of 0.22 lower, and the score for all of the questions was more uniform. This fact facilitates more significant positive correlations between these variables. However, there were statistically significant differences in the level of presence and satisfaction experienced by the users between the two groups in favor of the ARGroup. There were no statistically significant differences for the other variables. The correlation plot for the ARGroup (Fig. 10) helps in the identification of the variable that is most closely related to the Satisfaction variable, which is the Perceived Competence. Our argument for this relationship is that the more expert a user considers herself/himself to be after completing her/his experience, the greater their overall satisfaction. For the level of presence, the two variables that are most closely related are Usability and Interest/Enjoyment and, to a lesser extent, Perceived Competence. Our arguments for these relationships are that ease in learning and handling has a positive influence on the level of presence. The enjoyment when using the app and how interesting the app seems to users are two factors that also contribute positively to the level of presence.

With regard to the correlations among the three variables related to the three presence questionnaires ([22], [23], [24]), there were significant positive correlations among these three variables for the two groups (the ARGroup and the NoARGroup), as shown in Fig. 11. To our knowledge, this is the first study in which questions of the three presence questionnaires are used, and, moreover, correlations are found among them. This result indicates that the selection of the questions is appropriate for measuring the level of presence in our task. This selection could be used in other works to check whether or not the trend is similar.

With regard to Witmer's study [24], we also used four factors. However, the questions that were included in each of our factors were a subset of those used in Witmer's study. In our study (Fig. 12), there were the same number of significant positive correlations among the four factors in the two groups (the ARGroup and the NoARGroup). We found more significant positive correlations than Witmer [24]. Out of the four relationships found by Witmer, we coincide on three. As in Witmer's study, in our work, Involvement is strongly related to Adaptation/Immersion and Visual Fidelity. We also found a relationship between Adaptation/Immersion and Interface Quality, which, in our case, was stronger. Our explanation for the differences in the relationship of Visual Fidelity with the rest of the factors is that Sensory Fidelity was used in Witmer's study and it includes visual, auditory, and haptic items. In our case, the senses of audio and touch have not been considered. This could also explain the close relationship between Visual Fidelity and Interface Quality. The relationship between Involvement and Interface Quality can be explained by the relationship of the questions included in each factor in our study. Our argument for this relationship is that a higher Interface Quality has a positive influence on Involvement. To our knowledge, this is the first study in which the 4-factor model of Witmer has been used as a base for measuring presence using a mobile AR app. Moreover, several questionnaires ([22]–[27]) were used as a base to evaluate the users' subjective experience using a mobile AR app.

In this initial study, the app does not control when the user sees the virtual objects in the real environment in the adaptation phase. If these objects are not seen in this phase, the AR and NoAR conditions are the same. Thanks to this study, we solved this problem and we have incorporated this control in the app. Therefore, we are sure that the user has seen all of the virtual elements in the real environment. Now, when the users find a virtual object, they must touch it on the screen and a green sphere with a certain level of transparency envelops the virtual object. This sphere does not disappear during the entire phase.

We used the Tango SDK for the development of the app. There are other SDKs (ARKit, https://developer.apple.com/arkit, and Google ARCore, https://developers.google.com/ar) with similar characteristics. However, when we developed our app, ARCore and ARKit did not offer the same functionality as the Tango SDK. The Tango SDK included the functionality to identify flat horizontal and vertical surfaces. This functionality has already been incorporated in ARCore as of May 8, 2018.

It would be very interesting to explore the potential of other devices such as head-worn displays that can run SLAM-based apps (e.g., Microsoft HoloLens (https://www.microsoft.com/en-us/hololens) or Magic Leap (https://www.magicleap.com)).

VII. CONCLUSIONS

We have developed the first SLAM-based AR app to assess short-term spatial memory. Our app is an authoring tool that allows the evaluators to perform the assessment in any indoor environment, to add the objects that they require, and to change the objects from one session to the next. We carried out a study involving 55 participants. The participants were divided into two groups: the ARGroup (participants who learned the location of the virtual objects using AR) and the NoARGroup (participants who learned the location of the objects by looking at photographs). The results show that the performance outcomes in remembering objects and their location were statistically significantly greater for the participants in the ARGroup than for the participants in the NoARGroup. That is, our main contribution is that touring the augmented environment helped the participants to better remember the location of virtual objects added to the real scene compared to looking at photographs of the environment. This new contribution can be exploited for the development of tasks to assess or train spatial memory in a way similar to the way that spatial memory performs in everyday life.

This is the first study that we have carried out with this app, but many more studies can be done. In this paper, we compared two conditions of the app. After demonstrating that AR based on SLAM helps in the memory of the location of virtual objects added to the real scene, more studies can be carried out. In our case, we plan to carry out a comparison of our app with traditional neuropsychological tests and involving people without disabilities or mental dysfunctions. This new study would corroborate the hypothesis that the results for our app would reflect the spatial short-term memory ability of participants in the same way as traditional procedures. The corroboration of this hypothesis would also strengthen the contribution of this paper. Another study that we plan to carry out is to test our app with acquired brain damage patients. This type of studies would demonstrate the potential of our proposal for different collectives. Another variable to analyze is the environment used. In our study, we used a small-scale environment, a living room. A study in a more controlled area could also be conducted to rule out context-contingent potential interferences of (unknown or even known) stimuli that could influence test participants. Our app also works in large-scale environments (e.g., several floors of a building such as a university). In another study, the advantages and disadvantages between small-scale and large-scale environments could also be analyzed. Currently, our app stores data about errors committed in the evaluation phase and the duration of the phases. Other data that could be stored are the paths followed in the adaptation and evaluation phases. These paths could be analyzed to identify patterns of behavior between groups. Our task could also be adapted to other types of devices (e.g., Magic Leap).

ACKNOWLEDGMENT

The authors would like to thank all the people who participated in the study as well as the reviewers for their valuable comments.

REFERENCES

[1] T. Bailey and H. Durrant-Whyte, ‘‘Simultaneous localization and mapping (SLAM): Part II,’’ *IEEE Robot. Autom. Mag.*, vol. 13, no. 3, pp. 108–117, Sep. 2006.

[2] 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. Hum. Behav.*, vol. 54, pp. 414–424, Jan. 2016.

[3] A. O. Caffò *et al.*, ‘‘Intervention strategies for spatial orientation disorders in dementia: A selective review,’’ *Develop. Neurorehabil.*, vol. 17, no. 3, pp. 200–209, 2014.

[4] A. M. Barrett and T. Muzaffar, ‘‘Spatial cognitive rehabilitation and motor recovery after stroke,’’ *Current Opinion Neurol.*, vol. 27, no. 6, p. 653, 2014.

[5] 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, 2014.

[6] 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. Individual Differences*, vol. 50, pp. 11–22, Aug. 2016.

[7] U. Rehman and S. Cao, ‘‘Augmented reality-based indoor navigation using Google glass as a wearable head-mounted display,’’ in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2015, pp. 1452–1457.

[8] J. Polvi, T. Taketomi, G. Yamamoto, A. Dey, C. Sandor, and H. Kato, ‘‘SlidAR: A 3D positioning method for SLAM-based handheld augmented reality,’’ *Comput. Graph.*, vol. 55, pp. 33–43, Apr. 2016.

[9] J.-C. Piao and S.-D. Kim, ‘‘Adaptive monocular visual–inertial SLAM for real-time augmented reality applications in mobile devices,’’ *Sensors*, vol. 17, no. 11, p. 2567, Nov. 2017.

[10] R. Egodagamage and M. Tuceryan, ‘‘Distributed monocular visual SLAM as a basis for a collaborative augmented reality framework,’’ *Comput. Graph.*, vol. 71, pp. 113–123, Apr. 2018.

[11] L. Chen, W. Tang, N. W. John, T. R. Wan, and J. J. Zhang, ‘‘SLAM-based dense surface reconstruction in monocular minimally invasive surgery and its application to augmented reality,’’ *Comput. Methods Programs Biomed.*, vol. 158, pp. 135–146, May 2018.

[12] A. D. Baddeley, ‘‘Working memory,’’ *Science*, vol. 255, no. 5044, pp. 556–559, Jan. 1992.

[13] 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.

[14] S. Walkowiak, T. Foulsham, and A. F. Eardley, ‘‘Individual differences and personality correlates of navigational performance in the virtual route learning task,’’ *Comput. Hum. Behav.*, vol. 45, pp. 402–410, Apr. 2015.

[15] 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 Disorders*, vol. 16, no. 2, pp. 203–207, 2014.

[16] R. A. Ruddle and S. Lessels, ‘‘The benefits of using a walking interface to navigate virtual environments,’’ *ACM Trans. Comput.-Hum. Interact.*, vol. 16, no. 1, p. 5, 2009.

[17] 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.

[18] D. Rodríguez-Andrés, M. Méndez-López, M.-C. Juan, and E. Pérez-Hernández, ‘‘A virtual object-location task for children: Gender and videogame experience influence navigation; age impacts memory and completion time,’’ *Frontiers Psychol.*, vol. 9, p. 451, Apr. 2018.

[19] 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 *Proc. Int. Conf. Comput. Graph. Theory Appl.*, 2018, pp. 46–57.

[20] D. A. Bowman and R. P. McMahan, ‘‘Virtual reality: How much immersion is enough?’’ *Computer*, vol. 40, no. 7, pp. 36–43, Jul. 2007.

[21] 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.

[22] H. Regenbrecht and T. Schubert, ‘‘Measuring presence in augmented reality environments: Design and a first test of a questionnaire,’’ in *Proc. 5th Annu. Int. Workshop Presence*, 2002. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.581.2484&rep=rep1&type=pdf

[23] 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.

[24] B. G. Witmer, C. J. Jerome, and M. J. Singer, "The factor structure of the presence questionnaire," *Presence, Teleoperators Virtual Environ.*, vol. 14, no. 3, pp. 298–312, Jun. 2005.

[25] 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, U.K.: Taylor & Francis, 1996, pp. 189–194.

[26] *Intrinsic Motivation Inventory (IMI)*. Accessed: Jul. 2018. [Online]. Available: http://selfdeterminationtheory.org/intrinsic-motivation-inventory/

[27] A.-M. Calle-Bustos, M.-C. Juan, I. García-García, and F. Abad, "An augmented reality game to support therapeutic education for children with diabetes," *PLoS ONE*, vol. 12, no. 9, p. e0184645, Sep. 2017.

[28] R. P. C. Kessels, M. J. van Zandvoort, A. Postma, L. J. Kappelle, and E. H. F. de Haan, "The Corsi block-tapping task: Standardization and normative data," *Appl. Neuropsychol.*, vol. 7, no. 4, pp. 252–258, 2000.

[29] R. Cánovas, M. Espínola, L. Iribarne, and J. M. Cimadevilla, "A new virtual task to evaluate human place learning," *Behav. Brain Res.*, vol. 190, no. 1, pp. 112–118, Jun. 2008.

[30] E. A. Spieker, R. S. Astur, J. T. West, J. A. Griego, and L. M. Rowland, "Spatial memory deficits in a virtual reality eight-arm radial maze in schizophrenia," *Schizophrenia Res.*, vol. 135, nos. 1–3, pp. 84–89, Mar. 2012.

[31] P. J. Rosa, D. Morais, P. Gamito, J. Oliveira, and T. Saraiva, "The immersive virtual reality experience: A typology of users revealed through multiple correspondence analysis combined with cluster analysis technique," *Cyberpsychol. Behav. Soc. Netw.*, vol. 19, no. 3, pp. 209–216, Mar. 2016.

[32] S. E. Gathercole, S. J. Pickering, B. Ambridge, and H. Wearing, "The structure of working memory from 4 to 15 years of age," *Develop. Psychol.*, vol. 40, no. 2, pp. 177–190, Mar. 2004.

[33] J. R. Best and P. H. Miller, "A developmental perspective on executive function," *Child Develop.*, vol. 81, no. 6, pp. 1641–1660, 2010.

[34] J. Bacca, S. Baldiris, R. Fabregat, S. Graf, and D. Kinshuk, "Augmented reality trends in education: A systematic review of research and applications," *Educ. Technol. Soc.*, vol. 17, pp. 133–149, Oct. 2014.

[35] D. Furió, S. González-Gancedo, M.-C. Juan, I. Seguí, and N. Rando, "Evaluation of learning outcomes using an educational iPhone game vs. traditional game," *Comput. Educ.*, vol. 64, pp. 1–23, May 2013.

[36] D. Furió, S. González-Gancedo, M.-C. Juan, I. Seguí, and M. Costa, "The effects of the size and weight of a mobile device on an educational game," *Comput. Educ.*, vol. 64, pp. 24–41, May 2013.

FRANCISCO MUNOZ-MONTOYA received the bachelor's degree in computer engineering and the master's degree in artificial intelligence, pattern recognition, and digital imaging from the Universitat Politècnica de València. He is currently pursuing the Ph.D. degree in computer science.

M.-CARMEN JUAN is currently a Full Professor of computer science with the Universitat Politècnica de València, where she has taught, since 1996. She has participated in over 50 projects. She has been a principal investigator in 15 projects. She has authored over 40 papers in journals indexed in Journal Citation Reports and about the same number of papers published in conferences indexed in CORE. She has also authored over 150 works published in other conferences or journals. Her main research fields include computer graphics, augmented reality, spatial memory, advanced user interfaces and their applications to psychology and edutainment.

MAGDALENA MENDEZ-LOPEZ is currently an Associate Professor with the Department of Psychology and Sociology, University of Zaragoza, Spain. Her current research interests include spatial learning, spatial memory assessment, virtual reality and augmented reality, and their applications to psychology.

CAMINO FIDALGO received the Ph.D. degree from the Department of Psychology, University of Oviedo, in 2012. She is currently an Assistant Professor with the Department of Psychology and Sociology, University of Zaragoza, Spain. Her area of interest is application of information and communication technologies to psychology.

• • •