Journal on Multimodal User Interfaces (2022) 16:319–333
https://doi.org/10.1007/s12193-022-00392-4

ORIGINAL PAPER

A SLAM-based augmented reality app for the assessment of spatial short-term memory using visual and auditory stimuli

M.-Carmen Juan[1] · **Magdalena Mendez-Lopez**[2] · **Camino Fidalgo**[2] · **Ramon Molla**[1] · **Roberto Vivo**[1] · **David Paramo**[1]

Received: 7 July 2020 / Accepted: 3 June 2022 / Published online: 18 July 2022
© The Author(s), under exclusive licence to Springer Nature Switzerland AG 2022

Abstract

A SLAM-based Augmented Reality (AR) app has been designed, developed, and validated to assess spatial short-term memory. Our app can be used with visual and auditory stimuli and can run on mobile devices. It can be used in any indoor environment. The anchors and data of the app are persistently stored in the cloud. As an authoring tool, the type of stimulus, its number, and specific positions in the real environment can be customized for each session. A study involving 48 participants was carried out to analyze the performance outcomes comparing the location and remembering of stimuli in a real environment using visual versus auditory stimuli. The number of objects placed correctly was similar for the two different stimuli used. However, the group that used the auditory stimulus spent significantly more time completing the task and required significantly more attempts. The performance outcomes were independent of age and gender. For the auditory stimuli, correlations among all of the variables of the AR app and the variables of two other tasks (object-recall and map-pointing) were found. We also found that the greater the number of correctly placed auditory stimuli, the greater the perceived competence and the less mental effort required. The greater the number of errors, the less the perceived competence. Finally, the auditory stimuli are valid stimuli that may benefit the assessment of the memorization of spatial-auditory associations, but the memorization of spatial-visual associations is dominant, as our results suggest.

Keywords Augmented reality · SLAM · Spatial memory · Auditory stimuli · Visual stimuli · Assessment

1 Introduction

Spatial short-term memory can be defined as the limited ability that people have to store and remember representations of spatial stimuli for short periods of time [1]. These stimuli can be visual, auditory, or other types of stimuli [2]. Spatial memory is used to memorize relevant information, for instance, where we have left our belongings, or the route to find a place that was previously visited, among other examples [3]. Therefore, this type of memory allows spatial abilities and navigation and is involved in everyday tasks. This implies that if there is an impairment in spatial memory, the effects on daily life can be devastating. Tools that are related to spatial memory can be used for both assessment and training. For assessment, they can help in identifying difficulties for people's independence [4]. For training, they can be used in patients with orientation difficulties that are usually due to diseases or impairments, such as Alzheimer's disease [5], acquired brain injury [6], stroke [7], or healthy aging [8]. In real life, human spatial orientation is based on both visual and self-motion cues (vestibular and proprioceptive) [9]. Therefore, Virtual Reality (VR) and Augmented Reality (AR) are two technologies that can be exploited for the development of systems for the assessment of spatial memory. From our point of view, these two technologies are especially interesting when the user has to perform physical movement in the real environment. Physical displacement is important in spatial ability [10]. To navigate in the real world, people use information from different sources (mainly visual and body-based) as well as various complementary cognitive processes. Movement can be perceived through optic flow (visual) and vestibular information and proprioception (both of which are based on the body). People can use path integration (body-

✉ M.-Carmen Juan
mcarmen@dsic.upv.es

[1] Instituto Universitario de Automática e Informática Industrial, Universitat Politècnica de València, Valencia, Spain

[2] Departamento de Psicología y Sociología, Universidad de Zaragoza, Zaragoza, Spain

based) and the proximity of a landmark to determine where to travel within an environment. Therefore, with physical movement, tasks are similar to those of daily life and the environment can be explored in a natural and ecological way. The Simultaneous Localization and Mapping (SLAM) technique allows us to build or update a 3D map of an unknown environment and also keep track of the device used. SLAM has played a key role in the fields of autonomous mobile robots, driverless vehicles, and unmanned aerial vehicles. In recent years, special interest has been shown in SLAM for its use in augmented and virtual reality. For example, SLAM is the tracking technique behind ARCore (the platform we use) and ARKit. The use of SLAM in AR allows virtual objects to be placed on the scene without the need to physically add additional elements to it, such as image targets or specific 3D objects. In this work, we present an AR app based on SLAM that requires the physical movement of the user and can be used in any indoor environment (e.g., the patient's home, the therapist's office, or a hospital ward). Moreover, our app can use visual or auditory stimuli. Our study compares short-term spatial memory involving the visual and auditory senses. Memory for locations of visual and auditory stimuli depends on independent systems [11, 12]. Humans are visually oriented in their ability to recognize what they perceive. This fact determines the quality of the representations they make with each sensory modality. The mental representations of visual objects are processed in a detailed manner in the visual memory system whereas the auditory objects are processed in a conceptual manner in the auditory memory system [12]. This means that humans can store more items with more details when they are perceived in the visual modality than in the auditory modality [11, 12]. To our knowledge, this is the first work that uses AR for comparing visual and auditory stimuli for the assessment of short-term spatial memory.

2 Related work

Spatial memory is generally assessed using the sense of sight and paper and pencil tests [13, 14]. Computerized tools using VR or AR have been developed and provide some advantages over traditional testing. The use of VR or AR applications offers several advantages over paper and pencil tests. VR and AR applications enable acquiring and storing all of the data obtained during the task (e.g., successes and errors, number of errors, times, etc.). In applications of this type, the stimuli can be varied and can be controlled [15–17]. Another important advantage is the reduction in cost and time. For all of these reasons, the use of VR and AR has been extended for the study of human ability.

Very few works have explored the use of other types of devices for the assessment of spatial memory [18]. The added value of the work of Loachamín et al. [18] was the use of ambient electronic devices in the shape of a rabbit (Karotz) and the use of auditory stimuli for the assessment of spatial memory. Five devices were distributed in a 5 m^2 room with 30º of separation between each other. The task included a total of 45 auditory stimuli, divided into five levels, in which the number of sounds to remember increased. The participants had to listen to sounds and remember the device that made the sound and then remember the location of the sound source. The interaction was through gestures. The participants stood in front of the devices and raised both arms. A Microsoft Kinect sensor was used for gesture recognition. They compared the performance of 48 adults and 100 children using their system with paper-and-pencil neuropsychological tests. The system performance correlated with the performance of the neuropsychological tests. The number of correct levels correlated with four neuropsychological tests (memory of places, verbal working memory, discrimination of sounds from the environment, and clinical evaluation of spatial memory in everyday life). These correlations demonstrated the similarity of their task and traditional methods.

2.1 Virtual reality

VR has been used for the assessment of spatial memory in humans [19–22]. The first VR applications to assess spatial memory generally used a computer screen to visualize tasks and used a keyboard and mouse to navigate and change the user's perspective in the virtual environment. Users were sitting in front of these screens, without moving, and exploring a virtual environment [16, 17, 23]. Today, this possibility is still a valid alternative [24]. The use of headsets has also been considered for creating more immersive VR experiences [25]. Moreover, the latest headsets allow users without stereopsis to have 3D experiences (e.g. Oculus) [26]. Other works have incorporated physical activity to move in the virtual environment [27–29].

Rodríguez-Andrés et al. [28] analyzed the influence of movement on a VR task. They compared the performance of physical activity versus no physical activity. For the physical activity, they used a Wii balance board and a Wii mote. For the case of non-physical activity, they used a gamepad. For the physical activity, movement in the virtual environment was achieved by walking on the Wii balance board and the turns of the avatar were controlled with a steering wheel which contains the Wii mote. They used a large screen (120") for visualization. They carried out a study with 160 children. The results did not show significant differences between the two types of interaction. However, it should be noted that physical activity was performed using the Wii balance board and the user may not have had the feeling of walking in the real world.

Another VR task based on a maze also required physical activity [27]. A study (N = 89) was carried out to compare physical activity versus non-physical activity. A real bicycle was used for the physical activity interaction and a gamepad for the non-physical activity interaction. A VR HMD (Oculus Rift) was used for visualization. Their results indicated that the outcomes on the task were better for the non-physical activity interaction. For the physical activity interaction, the participants had to pedal on a stationary bike, use the handlebar, and wear the HMD. All of this could have influenced the usability of the system and the level of presence induced, leading to a negative influence on the results.

2.2 Augmented reality

To our knowledge, very few works have used AR to study the performance of spatial memory. In a first work, an AR app for visual stimuli using Vuforia SDK[1] was developed using image targets distributed in the real environment [15, 30]. The user had to physically walk around a room to look for objects that were distributed throughout the room inside of carboard boxes and remember their locations. Then, the user had to remember the position in which the objects were found. In a second work, a SLAM-based AR app for visual stimuli using Tango SDK was developed [31, 32]. The user also had to physically walk around a room and look for objects distributed in the room; however, in this case, there were no additional elements. These aforementioned applications only consider visual stimuli while this work incorporates multimodal interaction (visual and auditory stimuli). We also present a study comparing the two types of stimuli. Another disadvantage of the work in [31, 32] is that the app can only run on smartphones that support Tango SDK. In contrast, the app presented in this paper can run on different mobile devices.

Another work that is related to ours developed an AR system to memorize the location of objects and estimate distances [33]. Keil et al. [33] used HoloLens to display grids on the room floor. They carried out a study (N = 60) to check if the use of grids helped in these processes. The authors concluded that distance estimations were more accurate when a grid was shown, while location memory performance was worse when a grid was shown.

Other research groups have used AR as a navigation aid. Peleg et al. [34] presented a route-planning task for public transportation. They carried out a study (N = 44) comparing the performance of older and younger participants using a mobile AR app and a non-AR app. The mobile AR app augmented a wall map with the times that busses go through each station. The participants using the AR app completed the task requiring less time, but with higher error rates. Chu et al. [35] presented a mobile navigation service with AR. Their study (N = 49) compared the performance of the participants using the AR system and the maps. They concluded that AR navigation was better than the two map navigation modes that were used to determine the time and accuracy rate in completing the task. Rehman and Cao [36] developed an AR app for indoor navigation. They carried out a study (N = 39) comparing the navigation using Google Glass, a smartphone, and a paper map. The comparison of Google Glass and the smartphone did not offer statistical differences in terms of workload. The smartphone and Google Glass were better than the paper map in terms of less time required for execution and lower workload. The paper map showed better route retention.

3 A SLAM-based augmented reality app for the assessment of spatial memory

3.1 Design rationale

This subsection describes our design decisions as well as the solutions and available technologies. Our aim was to design an AR app for the assessment of spatial memory that is functional and effective. Our first step was to analyze the main functional characteristics. The app must include multimodal interaction. The interaction must be as natural and intuitive as possible. The app must include 3D objects that are familiar to users and that can appear in familiar environments. It must not require adding additional elements to the real world, i.e., it must work in any environment. It must be possible to configure tasks. These tasks must be stored so that they can be used by any user later. The app must persistently store the user's performance data when using the app. It must be possible to consult the user's performance data even when the application is not running.

We had to make several decisions regarding these characteristics. The first decision to make was to determine what kind of stimuli to include. We decided to include two different stimuli. The first stimulus selected was sight because it is the dominant sense in humans [37, 38]. For the other stimulus, two alternatives were studied: audio and touch. Touch was ruled out due to the difficulty of integrating the stimuli successfully in an AR app. The next decision was to program the app to appear in portrait mode to prevent the camera from being covered by the user's hand and to facilitate the manipulation of the device. A translucent button bar is added with buttons that show the objects to select. The translucence allows the user to see them, but the real world can also be seen in the background. The button bar can be placed on any of the four sides of the screen. It was suggested to place the bar at the bottom. This area is where menus usually appear and it is easy for both right-handed and left-handed people to

---

[1] https://developer.vuforia.com

manipulate. When placing an object on the scene, the user can use a manipulator to change the object's position, rotate it, or raise it through gestures on the screen. For simplicity and to be similar to the gestures already used in applications of this type, we use standard gestures. To select an object, the user only has to touch the area of the screen in which it appears. To rotate a selected object, two fingers are rotated (generally, the thumb and index fingers) with a separation between them of about 3 cm. Raising an object from its current position requires a two-finger swipe from the bottom of the screen to the top. This type of interaction is included natively in ARCore. We decided to select everyday objects that may be familiar to users and that can be included naturally in everyday environments. Moreover, the selected objects must have a 3D model and a specific sound that can be associated with it to prevent confusion with another 3D object. The objects to be considered are: mobile phones, mugs, cameras, musical instruments, and decorative objects (e.g., cars, animals, etc.). These 3D objects can be obtained from already available models, by accessing free or paid 3D models on the internet or modeling them from scratch (when required). We studied possible 3D design software to model or modify 3D objects: 3D Studio Max and Blender. Blender was selected because it is free and the group had already worked with it. For the development of AR apps, we had to decide between SDKs that require the inclusion of additional elements (Vuforia) or those that do not require the inclusion of such elements and are SLAM-based (Vuforia, Tango, ARCore, ARKit). To meet the requirements, we had to select an SDK that does not require the inclusion of additional elements. For this reason, the use of Vuforia with targets (image targets, model targets, or object targets) was ruled out. Similarly, the use of Vuforia Ground Plane was ruled out because it only detects horizontal surfaces. Tango was discarded because it has been deprecated since March 1st, 2018. ARKit was discarded since a Mac must be used for its programming and an iOS device for its execution. Consequently, we selected ARCore. Even though Unreal and Unity can both be used to program the apps, we selected Unity because the group had previously worked with this engine and switching to Unreal would not provide any additional benefits. A greater level of detail of the software and hardware used is included in the Appendix. A synthetic view of the architecture of the app is shown in Fig. 1. To be able to store where objects are placed so that they can be retrieved later, our app must have the functionality to persistently store anchors that are placed at a given position. After an exhaustive search, we determined that the only option available at that time for use with ARCore was Azure Spatial Anchors. For the storage of user data, several alternatives were studied: store the data on the mobile device, a server, or the cloud. We decided to store data in the cloud. In this way, the data is synchronized in real time with the app and is available for consultation even when the app is offline. After studying several options, we chose Firebase, which is a platform that is located in the cloud and integrated with Google Cloud Platform. Firebase was also selected because it allows easy management of user access to the app through email and other platforms such as Google, Twitter, or Facebook.

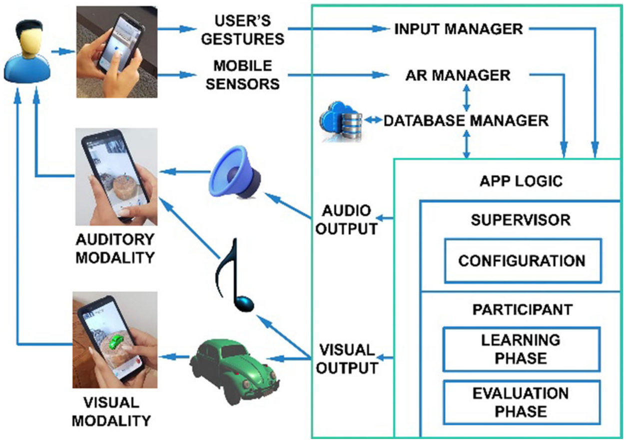


Fig. 1 A synthetic view of the architecture of the app

The app was developed from this design and can work on any device that supports ARCore. The app works with visual and auditory stimuli and can work in any indoor environment. The person in charge of the evaluation (hereinafter, supervisor) only has to customize the environment with the type of stimuli (visual or auditory), the number of stimuli, and the stimuli placed in the desired positions. The app can store different environments and their customization so that the supervisor only has to select one of them for a specific session.

3.2 App functionality for the supervisor

The user logs in with either a Google account or an e-mail account. After log-in, a screen with three buttons (placement, learning, and evaluation) appears with a configuration option in the lower area of the screen. By selecting the 'Configuration' option, the supervisor can define the maximum time for the phases to be performed by the participants. By default, there is no maximum time for these phases. By using the 'Placement' button, the supervisor can select the number of stimuli to be used in a session. Images of the different stimuli are shown on the screen for this selection. The eight stimuli included in this version of the app are: a camera, a car, a bell, a frog, a dog, a mug, a cell phone, and a violin. Each image has both a visual and an auditory stimulus associated to it. The stimuli were selected so that they had a visual stimulus and its sound equivalent. For example, the image of the dog has a 3D model of a dalmatian dog for the visual stimulus and a dog's bark for the auditory stimulus. All auditory stimuli

Fig. 2 Example of the AR app. A surface is detected and appears as a plane with white circles

are placed in the environment using a musical note. Therefore, for the selection and placement of the objects in the real environment, the supervisor only uses visual stimuli. When a user performs a session, either visual or auditory stimuli are selected for use by the supervisor. After the selection of the stimuli, the supervisor must place them in the desired place in the real environment. To do this, the app has to identify flat surfaces. Planes with white dots are used to show the detected flat surfaces. The supervisor has to move the device around the real environment so that the app can identify suitable flat surfaces. The supervisor can place the stimuli in the environment on the suitable surfaces that appear as white circles (Fig. 2). The stimuli to be placed appear in the right or bottom sidebar of the screen. The bar on the right is landscape mode and the bar at the bottom is vertical mode. The supervisor chooses one of the objects and places it in the environment by tapping on the selected location. After placement, each object has a manipulator for changing its position (by using drag and drop), rotating it, or enlarging or reducing it on the screen. The selected object can be rotated by rotating two fingers on the screen. The size of the object can be enlarged or reduced by spreading two fingers apart or pinching two fingers together on the screen, respectively. The information about the environment, the stimuli, and their placement are persistently stored with an identifier that can later be recovered for another session.

3.3 App functionality for the user

In the current version of the app, the user must perform two phases consecutively (learning and evaluation). These phases are executed when selecting the ‘Learning’ and ‘Evaluation’ buttons of the main menu. In the learning phase, the type of stimulus to be used (visual or auditory) must be selected first. By default, it is visual. The supervisor then selects the environment desired from the stored environments. Afterwards, the user initiates her/his experience. The user walks around the real environment and looks at it through the screen of the mobile device. When the app detects an area where a stimulus has been placed, it shows it on the screen. Planes with white circles to indicate detected surfaces do not appear in this phase. Figures 3a and b show two screenshots of the app using the visual and auditory stimuli, respectively. These are two examples of the initial appearance of the objects before the user marks them as seen/heard. In the auditory version, when a 3D model of a musical note is touched, the associated sound is reproduced for three seconds. All the user has to do is tap on the screen in the area where the object is. The musical note appears about 20 cm. away from the flat surface on which the stimuli has been placed. In this auditory version, when the user taps on a note, its color changes to green in order to indicate that this sound has already been heard. In the visual version, once an object has been tapped, its color also changes to green (Fig. 4). This tap ensures that the user has seen the object. Figure 5 shows an environment that was prepared for this study in which eight objects can be observed in the same scene all together. Once all of the stimuli to search for have been marked as seen, a confirmation button appears. Depending on the settings defined by the supervisor, the user may have a limited amount of time in which the stimuli are visible.

In the evaluation phase, the objects to place appear in the right/bottom sidebar. None of the objects are shown in the real environment. Planes with white circles to indicate detected

(a)

(b)

Fig. 3 Examples of the AR app. The virtual objects are shown in their original colors because the user has not touched any yet. **a** View of the app with 8 visual stimuli. **b** View of the app with 8 auditory stimuli presented as blue music notes

Fig. 4 Example of the AR app with visual stimuli. View of the app with 8 visual stimuli after all being touched by the user shown in green

Table 1 Distribution of participants by gender and age

Years old	Men	Women
10–15	5	2
16–20	0	0
21–30	7	7
31–40	0	2
41–50	5	6
51–60	3	3
61–70	2	2
> 70	3	1

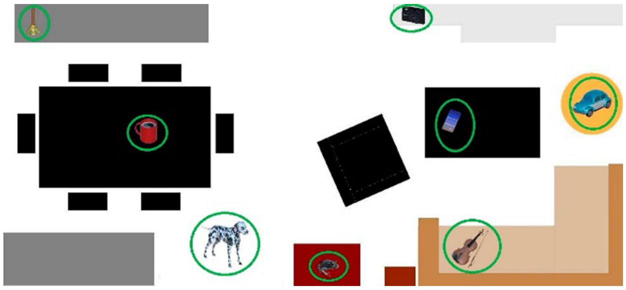


Fig. 5 Distribution of the 8 stimuli in the room in which the study was carried out. The stimuli are highlighted with green ovals

surfaces do not appear in this phase. The user chooses an object to be placed in the environment by tapping on it in the sidebar. After this selection, the user moves the mobile device and with the camera locates the place where she/he thinks she/he remembers the object was and taps on this location. If the object has been placed in the correct position, the object appears in that position. The objects do not have to be placed in exactly the 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 margin of error is the same for visual and auditory stimuli. If the object has not been placed in the correct position, the user can make two more attempts. If on the third attempt the object has not been placed in the correct position, the object is shown in that position, but the app stores these as failed attempts. The app stores all of the user's data regarding attempts and successes/failures.

4 User study

4.1 Participants, procedure, and measures

A total of 48 users participated in the study, of which 23 were women (48%), ranging in age between 11 and 87 years old (39.75 ± 19.9). Table 1 shows the distribution of participants by gender and age. This distribution can also be seen in Fig. 8. The study was approved by the Ethics Committee of the Universitat Politècnica de València, Spain. The study was conducted in accordance with the declaration of Helsinki.

The participants were divided into two groups: VisualGroup and AuditoryGroup. The participants in VisualGroup used the app with visual stimuli first, and one or two days later, they used the app with auditory stimuli. The participants in AuditoryGroup used the app with auditory stimuli first, and one or two days later, they used the app with visual stimuli. The reason for the second use of the app by the participants was to compare the two types of stimuli and to know the participants' opinion. The participants used the app twice (with auditory and visual stimuli) in order to be able to compare them and to indicate their preference. One or two days (depending on participants' availability) were left between the two sessions to avoid fatigue in some of the participants (14 of 48 participants were older than 60 years old).

Both groups were balanced in such a way that there were 23 users in VisualGroup (48%) and 25 users in AuditoryGroup (52%). The proportion of women was similar for both groups. The sessions with users were carried out in a room that had an area of around 30 m^2 with furniture that is commonly found in a living room. Figure 5 shows the distribution of stimuli in the room. Figure 6 shows an example of the app running on the room used in the study.

The protocol used in the study was the following. First, the supervisor explains what the task consists of to the participant, giving basic explanations about the app. The learning phase begins, in which the participant must find all of the virtual stimuli (visual or auditory) distributed throughout the room. To ensure that the participant does not forget to see/hear any of the stimuli, the objects or audios (represented by a musical note) to be found are shown at the bottom or on the right side of the screen, strictly for information purposes. The participant must tap on the objects or musical notes that are mixed in the environment in order to confirm that she/he has seen/heard the stimuli, and the color of the objects or musical notes changes to green. Once the participant has tapped all of the stimuli, a green confirmation button that returns to the main menu appears. The total amount of time of the learning phase and the time in which the user has

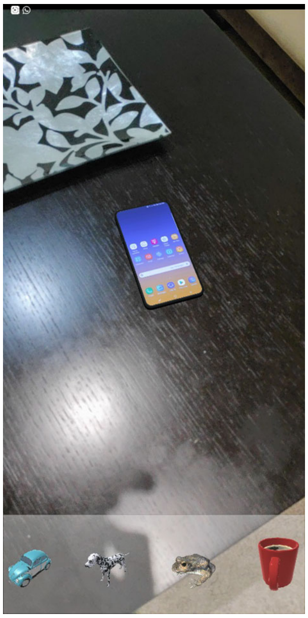


Fig. 6 An example of the AR app in which the virtual mobile phone is shown on the real table of the room used in the study

touched each stimuli is recorded during this phase. In our study, there was no time limit for memorizing the stimuli.

In the evaluation phase, the participant must select a stimulus (object or sound) from the right/bottom bar and tap on the screen in the location where she/he thinks the stimulus should be correctly placed. The participant has three attempts to place each of the stimuli. If the participant fails in the three attempts, the stimuli is placed in the position that was last touched.

The participant is then asked which stimuli she/he remembers and the supervisor writes them down. The participant is also asked to drawn each stimuli she/he remembers on the map in its correct location according to the AR task. Finally, the participant is asked to fill out Questionnaire 1 in order to determine her/his user experience with the app. As mentioned above, after one or two days, the participants repeated the process in order to perform the task with the other stimulus. After completing the task, the participants filled out a comparative questionnaire (Questionnaire 2) of the two experiences using the two different stimuli.

The Questionnaire 1 consists of 22 questions that we group in the following variables: enjoyment, concentration, usability, competence, calmness, expertise, mental effort, physical effort, satisfaction, and presence. This questionnaire was designed specifically for this study and was based on previously used questionnaires [31, 39–43].

The variables from the app used in the analysis are the following: (1) total number of stimuli found correctly in the task using visual or auditory stimuli (LocStimuli); (2) total number of errors made to correctly place a stimulus in the task using visual or auditory stimuli (Errors); (3) total time spent to perform the task using visual or auditory stimuli (Time).

Two additional tasks that were performed after using the AR app were also included in our study. The first one was the object-recall task, which consisted of free recall of the eight stimuli examined using our AR app. The supervisor asked participants "Which stimuli do you remember examining when using the AR app?” The participant verbally indicated the stimuli and the supervisor wrote them down. No feedback was given on responses. The variable used in the analysis is the total number of stimuli remembered correctly after the visual or auditory task (Recall). The second one was the map-pointing task, which tested the ability of the participants to read a bi-dimensional map of the room in which the AR app with the visual/auditory stimuli was used. The variable used in the analysis is the number of stimuli that the participant placed correctly in the map (Map).

5 Results

The normality of the data was verified using the Shapiro–Wilk test [44]. The tests indicated that the sample does not fit a normal distribution. Therefore, we used non-parametric tests (the Mann–Whitney U test, the Kruskal–Wallis test, and the Spearman correlation). A descriptor of each group is presented in the format (median (*Mdn*); interquartile range (*IQR*)). All of the tests are presented in the format (statistic *U/W*, normal approximation *Z*, *p-value*, *r* effect size). A statistically significant difference at level $\alpha = 0.05$ is indicated by the symbol **. The R open source statistical toolkit[2] was used to analyze the data (specifically, R version 3.6.2 and RStudio 1.2.5033 for Windows).

5.1 Performance outcomes

To determine how the use of visual or auditory stimuli affects memory of the location of the different stimuli when using the app, we compare the performance results between the two groups (VisualGroup vs. AuditoryGroup) in their first contact with the app (between-subjects analysis). First, we consider the variable that indicates the total number of successes for the eight stimuli used (LocStimuli). To determine whether or not there were differences in placing stimuli in their correct locations between the participants of VisualGroup ($Mdn = 8$; $IQR = 1.5$) and AuditoryGroup ($Mdn = 7$; $IQR = 3$), we applied the Mann–Whitney U test ($U = 351$, $Z = 1.434$, $p = 0.155$, $r = 0.207$). Figure 7 shows boxplots of LocStimuli variable and variables related with performance

---

[2] https://www.r-project.org

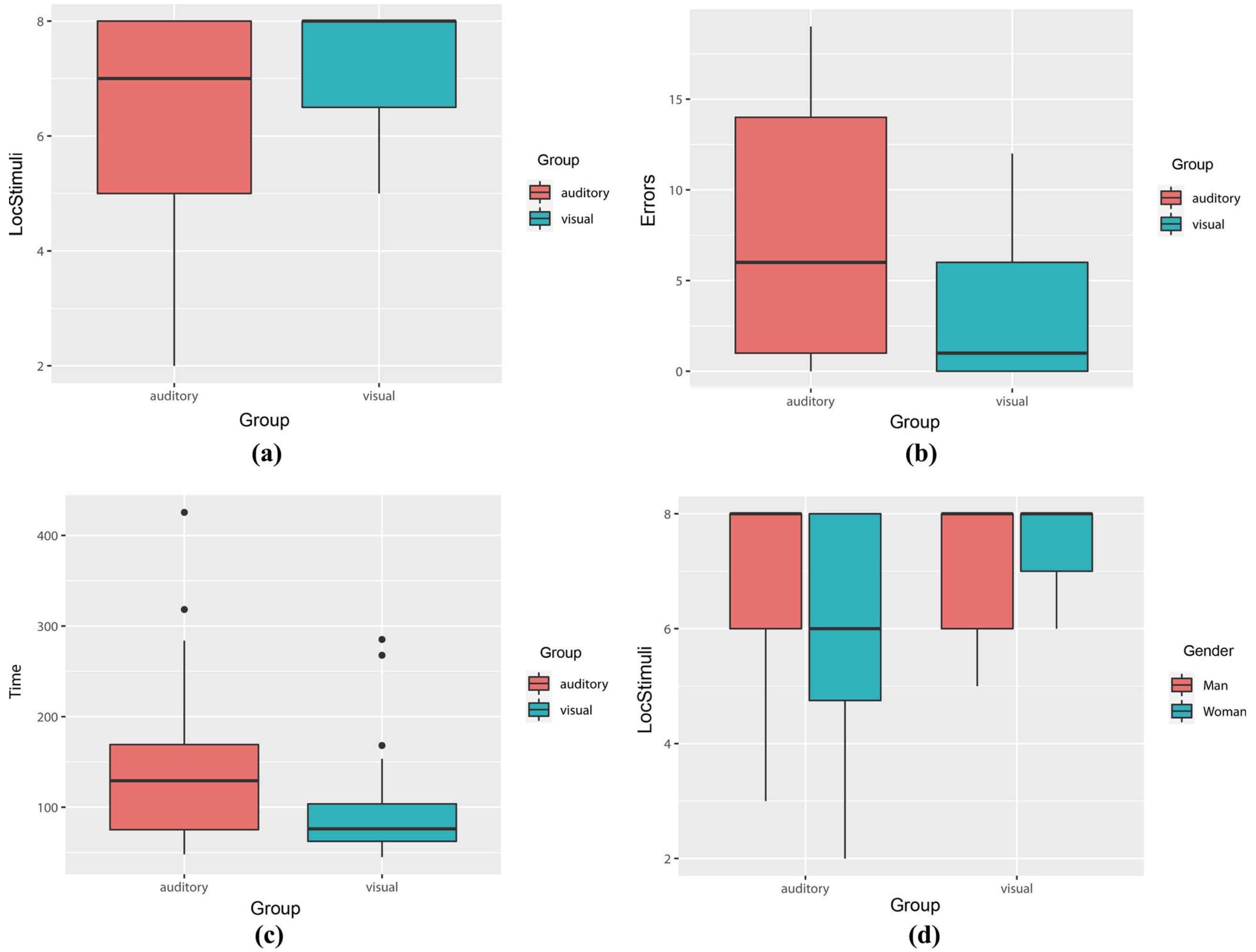


Fig. 7 Boxplots for variables related to the performance outcomes considering group and gender as factors. **a** Total number of stimuli correctly placed by group. **b** Total number of errors by group. **c** Total time required in seconds for the evaluation phase by group. **d** Total number of stimuli correctly placed by group and gender

outcomes considering the group and gender as factors. This result indicates that there are no statistically significant differences between the two groups.

Second, we consider the variable that indicates the total number of errors (Errors) made when placing the stimuli in the correct position. To determine whether or not there were differences for this variable between the participants of VisualGroup (*Mdn* = 1; *IQR* = 6) and AuditoryGroup (*Mdn* = 6; *IQR* = 13), we applied the Mann–Whitney U test ($U = 170.5$, $Z = -2.452$, $p = 0.015^{**}$, $r = 0.354$). This result indicates that there are significant differences between the two groups in favor of the group with visual stimuli, which required fewer attempts to place the stimuli correctly.

Third, we consider the variable that indicates the total time in seconds used to perform the evaluation phase (Time). To determine whether or not there were differences for this variable between the participants of VisualGroup (*Mdn* = 76.15; *IQR* = 41.33) and AuditoryGroup (*Mdn* = 129.24; *IQR* = 93.79), we applied the Mann–Whitney U test ($U = 184$, $Z = -2.136$, $p = 0.033^{**}$, $r = 0.308$). This result indicates that there are significant differences between the two groups in favor of the group with visual stimuli, which spent less time completing the test.

The variable that indicates the number of stimuli that the participant remembers verbally after performing the test with the app was also considered (Recall). To determine whether there were differences for this variable between the participants of VisualGroup (*Mdn* = 8; *IQR* = 0) and AuditoryGroup (*Mdn* = 7; *IQR* = 2), we applied the Mann–Whitney U test ($U = 402.5$, $Z = 2.837$, $p = 0.005^{**}$, $r = 0.409$). The variable that represents the number of stimuli placed correctly in the map was also analyzed (Map). To determine whether there were differences for this variable between the participants of VisualGroup (*Mdn* = 8; *IQR* = 0) and AuditoryGroup ($Mdn = 8$; $IQR = 4$), we applied the Mann–Whitney U test ($U = 384.5$, $Z = 2.344$, $p = 0.020^{**}$, $r = 0.338$). These results indicate that there are significant differences between the two groups in favor of the group with visual stimuli, which recalled more stimuli and placed more stimuli correctly on the map.

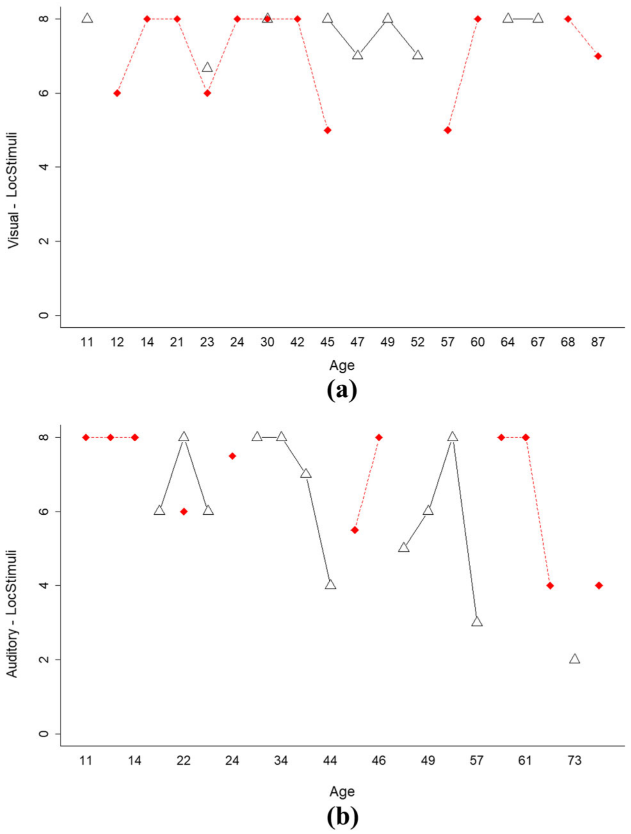
5.2 Gender and age analysis

To determine if gender influences the LocStimuli variable for VisualGroup, we applied the Mann–Whitney U test ($U = 74.5$, $Z = 0.597$, $p = 0.574$, $r = 0.124$); for AuditoryGroup, we also applied the Mann–Whitney U test ($U = 56.5$, $Z = -1.244$, $p = 0.224$, $r = 0.249$). To take into account gender and the Time variable, the same test was given to VisualGroup ($U = 61$, $Z = -0.308$, $p = 0.786$, $r = 0.064$) and to AuditoryGroup ($U = 91$, $Z = 0.707$, $p = 0.503$, $r = 0.141$). To take into account gender and the Recall variable, the same test was given to VisualGroup ($U = 72$, $Z = 0.631$, $p = 0.563$, $r = 0.132$) and to AuditoryGroup ($U = 66$, $Z = -0.697$, $p = 0.504$, $r = 0.139$). Similarly, to take into account gender and the Map variable, the same test was given to VisualGroup ($U = 59.5$, $Z = -0.555$, $p = 0.609$, $r = 0.116$) and to AuditoryGroup ($U = 68.5$, $Z = -0.558$, $p = 0.597$, $r = 0.112$). No statistically significant differences were found in any of these analyses, and, therefore, we can conclude that the performance results were independent of the participants' gender.

To determine if age influences the LocStimuli variable, we applied the Kruskal–Wallis test to VisualGroup ($\chi^2(17) = 15.757$, $p < 0.541$) and to AuditoryGroup ($\chi^2(21) = 19.426$, $p < 0.558$). Similarly, to take into account age and the Time variable, the same test was given to VisualGroup ($\chi^2(17) = 13.179$, $p < 0.724$) and to AuditoryGroup ($\chi^2(21) = 23.114$, $p < 0.338$). To take into account age and the Recall variable, the same test was given to VisualGroup ($\chi^2(17) = 22.0$, $p < 0.185$) and to AuditoryGroup ($\chi^2(21) = 22.48$, $p < 0.372$). To take into account age and the Map variable, the same test was given to VisualGroup ($\chi^2(17) = 22.0$, $p < 0.185$) and to AuditoryGroup ($\chi^2(21) = 23.31$, $p < 0.328$). No statistically significant differences were found in any of these analyses, and, therefore, we can conclude that the performance results were independent of the participants' age.

These analyses were performed using the whole sample. Since the sample is composed of different age ranges (from 11-year-olds to 87-year-olds), Fig. 8 shows two interaction plots for the LocStimuli variable (auditory and visual), considering the age and gender of the participants.

Figure 8a shows that, in the group of participants who used the app with the visual stimuli, there are no great differences for gender and age. Figure 8a also shows that, for visual stimuli, older adults do not present a great dispersion for the LocStimuli variable. Older adult dispersions during performance could reduce the possibilities of finding gender differences, which is not the case for the visual stimuli and the LocStimuli variable.



(a)

(b)

Fig. 8 Interaction plots for the total number of correctly placed stimuli, considering the age and gender of the participants. The red rhombuses indicate men. **a** Group that used the app with the visual stimuli first. **b** Group that used the app with the auditory stimuli first

Figure 8b shows that, in the group of participants who used the application with the auditory stimuli, there are more differences than for the visual stimuli for age, but not for gender. Participants under 44 years old correctly placed more stimuli. To determine whether or not there were differences in placing auditory stimuli in their correct locations between participants with an age less than 44 years old ($Mdn = 8$; $IQR = 1.25$) and greater than or equal to 44 years old ($Mdn = 5$; $IQR = 4.0$), we applied the Mann–Whitney U test ($U = 112$, $Z = 1.967$, $p = 0.053$, $r = 0.393$). It indicated that there are no statistically significant differences. Figure 8b also shows that, for auditory stimuli, adults over 55 years old present more dispersion for the LocStimuli variable. To determine whether or not there were differences in placing auditory stimuli in their correct locations between participants with an age less than or equal to 55 years old ($Mdn = 8$; $IQR = 2.0$) and older than 55 years old ($Mdn = 4$; $IQR = 3.75$), we applied the Mann–Whitney U test ($U = 82.5$, $Z = 1.726$, $p =$

Table 2 Significant Spearman correlations for AuditoryGroup (N = 25). A correlation of .1 ≤|r|<.3 has a small effect, of .3 ≤|r|<.5 has a medium effect, and of |r|≥ .5 has a large effect. $^{*}p < .05$; $^{**}\ p \leq .001$

	2	3	4	5
1. Errors	.53*	– .92**	– .52*	– .70**
2. Time		– .47*	– .74**	– .82**
3. LocStimuli			.50*	.65*
4. Recall				.76**
5. Map				1.0

0.091, $r = 0.345$). It indicated that there are no statistically significant differences.

The same analyses were performed for the Errors and Time variables. No statistically significant differences were found for any of these analyses, except for the Time variable when the auditory stimuli were used, in favor of the younger participants. However, as can be seen in Fig. 7, there are two outliers, corresponding to two men, one 72 years old and one 75 years old. If these two outliers are eliminated and the analysis is repeated, no statistically significant differences are observed. Therefore, we can conclude that the performance results were independent of the participants' gender and age.

5.3 Correlations

To calculate the correlation between the variables obtained in the app and the object-recall and map-pointing tasks, we used Spearman's correlation. For VisualGroup, a significant correlation was found between the Time variable and the Recall variable, with a correlation of − 0.54, $p < 0.01^{**}$. Table 2 shows the significant Spearman correlations for AuditoryGroup.

5.4 Subjective perceptions

The questionnaire about subjective perception was used to measure the participants' subjective perception about the AR app and their performance in the VisualGroup and AuditoryGroup. The questions were grouped in the following variables: enjoyment, concentration, usability, competence, calmness, expertise, mental effort, physical effort, satisfaction and presence. For VisualGroup and considering gender, only a significant difference was found for calmness ($U = 93.5$, $Z = 2.350$, $p = 0.021^{**}$, $r = 0.490$) in favour of the women. For AuditoryGroup and considering gender ($U \geq 57.5$, $Z \geq -1.206$, $p \geq 0.239$), no significant differences were found. Therefore, we can conclude that the subjective perception was practically independent of the participants' gender.

We used the Spearman rank correlation to test the associations among the subjective perception and the variables of the performance outcomes of the participants for the AR spatial task for the two groups. For VisualGroup, the only marginal correlation found was the Time variable with presence ($r = 0.41$, $p = 0.051$). For the correlations between the subjective variables, enjoyment correlates with concentration ($r = 0.63$, $p = 0.001$), competence ($r = 0.42$, $p = 0.049$), and calmness ($r = 0.52$, $p = 0.012$). Concentration correlates with perceived competence ($r = 0.84$, $p < 0.001$) and calmness ($r = 0.62$, $p = 0.002$). Usability marginally correlates with calmness ($r = 0.40$, $p = 0.056$). Perceived competence correlates with calmness ($r = 0.64$, $p = 0.001$) and non-mental effort ($r = 0.54$, $p = 0.008$). Table 3 shows the correlations for AuditoryGroup.

5.5 Preferences and open questions

After completing the two tasks, when the participants were asked "Which one did you like the most?", 79.2% of the participants preferred the app with the visual stimuli over the auditory stimuli. The main comment was that the task was more entertaining since they could see the augmented objects, or simply because they found it easier and ended up with a greater sense of satisfaction. The participants who preferred the app with the auditory stimuli did so because they found it to be a greater challenge, which made it more stimulating.

When asked which of the two apps was better for remembering the location of stimuli, practically all of the participants (89.6%) indicated that the app with the visual stimuli seemed better. The main argument was that it was easier to remember the positions of the stimuli because an association between the stimulus and the object that represents it could be established.

For the open question, "What do you think this technology could be used for?" Some of the comments included: (1) for exercises with children to learn objects and shapes; (2) to strengthen memory; (3) for space design, AR games, storyboard planning for cinema; (4) to help people with Alzheimer's disease; (5) to exercise memory and concentration; (6) to help children with problems of concentration and association of objects and sounds; (7) for shows or entertainment.

6 Discussion

AR technology provides greater ecological validity since the tasks to be carried out in the augmented environment simulate real situations, i.e., situations that are similar to what a user could do at home to locate personal belongings. In addition, the user has to physically move through the real world

Table 3 Significant Spearman Correlations for AuditoryGroup (N = 25). A correlation of $.1 \leq |r| < .3$ has a small effect, of $.3 \leq |r| < .5$ has a medium effect, and of $|r| \geq .5$ has a large effect. $^{\#}$ marginal, * $p < .05$; ** $p < .01$

	2	3	4	5	6	7	8	9	10	11	12
1. LocStimuli	− .92**	.33	− .13	− .14	.53**	.05	.08	.37$^{\#}$	− .23	.22	− .21
2. Errors		− .21	.02	− .02	− .58**	− .20	− .16	− .37$^{\#}$	.12	− .16	.18
3. Enjoyment			.14	− .06	.43*	.03	.00	.35	.31	.72**	.16
4. Concentration				.74**	.43*	.44*	.57**	.46*	.53**	.30	.76**
5. Usability					.26	.49*	.68**	.21	.49*	.15	.49*
6. Competence						.31	.09	.67**	.05	.30	.15
7. Calmness							.50*	.44*	.33	.25	.33
8. Expertise								.36$^{\#}$	.39$^{\#}$	.31	.49*
9. No-mental effort									.20	.35	.33
10. Non-physical effort										.45*	.58**
11. Satisfaction											.48*
12. Presence											1.0

to navigate in the augmented environment, just as the user would do at home when searching for his/her belongings. Therefore, these types of applications are especially suitable for spatial memory tasks (assessment or training).

Our AR app uses a multimodal interaction (visual and auditory) to investigate cognition. The idea of using auditory stimuli that are distributed in certain locations in a room is not new [18]. However, our proposal is dynamic. In the proposal of Loachamin et al. [18], the devices were fixed and had to be physically moved to be able to use them in another room or with another configuration. Our proposal is totally configurable and the supervisor can customize the sounds to be used and can place them in any environment. Another little explored sense is that of touch [45]. If these senses were used individually or in combination they could be of great help for spatial memory tasks.

Our SLAM-based AR app can work with any mobile device that supports ARCore. This versatility in the mobile device makes it much more accessible than previous apps, which could only work on mobile devices with support for Tango, for which there is currently none on sale in the market [31]. To improve accessibility, the development could be migrated to ARFoundation[3]. ARFoundation allows working in Unity with different AR platforms. For this, the platforms to be used must be integrated into Unity (ARKit XR Plugin for iOS or ARCore XR Plugin for Android). By integrating the two platforms, the app runs on devices with support for ARCore and ARKit.

The use of Azure Spatial Anchors allowed our app to store augmented reality anchors persistently in the cloud. The Azure Spatial Anchors stored in the cloud are accessible for HoloLens, iOS, and Android devices. If Azure Spatial Anchors had not been used, the supervisor would have had to physically place the stimuli in each session. The use of Firebase greatly facilitated the management of the supervisor's access to the app. Moreover, the data were stored in the cloud.

There was no statistically significant difference for the stimuli placed correctly between the two groups using the two different stimuli. However, the AuditoryGroup spent significantly more time completing the task and required significantly more attempts. Moreover, the VisualGroup was able to verbally recall a significantly higher number of stimuli and placed more stimuli correctly on the map after performing the AR task. From these results, we can conclude that AR apps that are in line with the one presented in this paper can be used as tools to assess spatial memory involving visual and auditory stimuli, and that visual stimuli positively affects short-term spatial memory outcomes. These results are in line with previous works stating that the sense of sight is the dominant sense in humans and, therefore, the learning of spatial-visual associations is facilitated [37, 38].

Our results show that the performance outcomes were independent of the age and gender of the participants. This demonstrates that, regardless of age and gender, our app has proven to be suitable for the assessment of spatial memory. These results are in line with previous works [15]. The age range of our study was between 11 and 87 years old (39.75 ± 19.9), and, in another study [15], the age range was between 8 and 72 years old (36.53 ± 15.78). When both age ranges are combined, this suggests that apps of this type are suitable for all ages (8 to 87 years old). However, cognitive functions decline with age. These decreases are associated with changes in patterns of functional activity [46]. Therefore, in other studies, researchers could determine if the task is sensitive enough to age and study more in depth the suit-

---

[3] https://docs.unity3d.com/Packages/com.unity.xr.arfoundation@3.0

ability of the app in its current state for older people or personalize it for that group.

The following correlations were identified regarding the correlations among the variables from the AR app (LocStimuli, Time and Errors) and the variables from the object-recall (Recall) and map-pointing (Map) tasks. For AuditoryGroup, significant correlations were found among all of them (Fig. 5). The result between LocStimuli and Map shows that the spatial-auditory associations were learned and transferred from the three-dimensional array of the real room to the bi-dimensional array of the map. On the other hand, this correlation was not found for VisualGroup. This result is contradictory with the results obtained in previous works, in which the memory for the location of virtual objects in visual modality in a navigational space correlated positively with the pointing of these objects on a map [30]. Our explanation for our non-correlation is that the LocStimuli scores were very high (with a mean of 7.3), the scores for the Map variable were also very high (with a mean of 7.6), and minimal variations in participant's scores influenced the non-significant correlation. Therefore, despite not finding such a correlation and after reviewing the scores, we can suggest that there was learning and transfer. Similar results were obtained between the LocStimuli and Recall variables. Our explanation for those results is the same as the one above.

With regard to the subjective variables, the results indicate that the app was highly appreciated by the participants as a whole using both types of stimuli in all of the variables analyzed. On a scale of 1 to 7, the medians were very high, equal to, or above 6 in all cases. The subjective perception was practically independent of the participants' gender. Only a significant difference was found for the Calmness variable and for VisualGroup, in favour of the women. The correlations among the variables related to the performance with the AR app and the subjective variables were analyzed for VisualGroup and AuditoryGroup. For VisualGroup, nine correlations were found. For AuditoryGroup, twenty-seven correlation were found. It is worth highlighting the following: the greater the number of correctly placed stimuli, the greater the perceived competence, the less mental effort required; the greater the number of errors, the less the perceived competence, and the less mental effort required; the more enjoyment experienced, the more the perceived competence, and the more satisfaction felt; the higher degree of usability, the calmer the participant, the more expertise felt, the less physical effort required, and the more sense of presence felt.

With regard to the question "Which one did you like best?", 79.2% of the participants preferred the app with the visual stimuli over the auditory stimuli. For the question "Which of the two apps seems best to remember the location of stimuli?", 89.6% of the users indicated that the app the with visual stimuli seemed better. At this point, it should be noted that the users who participated in our study had no vision problems, and, therefore, their dominant sense is that of sight. Thus, it is understandable that they preferred the sense of sight. Moreover, with the visual stimulus, the participants saw 3D objects that were mixed in with the real environment, while with the auditory stimulus, they only saw a musical note, which was always the same. Also, since visual stimuli in AR attract users, they focus on them more and this benefits their memory [31]. Therefore, a future study could focus on testing our auditory system with people who have visual impairments.

Several studies have identified spatial orientation deficits in patients with brain injury [47]. People suffering from a neurological impairment can also misplace objects [48]. In an ongoing study, we are comparing patients with acquired brain injury with healthy participants in a task involving visual stimuli. We have observed that the patients are able to perform the task with the app and that they appreciate and accept the AR app. Therefore, from this preliminary observation, we consider that our app could be a useful tool for those and other groups with problems related to spatial memory.

7 Conclusion

We have described the development and validation of the first SLAM-based AR app for the assessment of short-term spatial memory that can work with any mobile device that supports ARCore. As an authoring tool, the supervisor can use our app in any indoor environment. She/he can select the type of stimuli to be used (visual or auditory), the number of stimuli, and the specific stimuli. The configurations of the environments are stored in the cloud and can be recovered when desired.

The combination of auditory stimuli and AR has not been used for the assessment of short-term spatial memory in a navigational space. In this work, we compared visual and auditory stimuli. From our results, we can conclude that our AR app can be used to assess spatial memory involving visual and auditory stimuli. The number of correctly placed stimuli was similar for both types of stimuli. The group that used the auditory stimuli required more time and more attempts to complete the task. Performance outcomes were independent of age and gender. Therefore, the auditory stimuli are stimuli with great potential that can help in assessing the ability to memorize spatial-auditory associations.

As future work, many studies can be carried out. A very interesting study could investigate the possible impact of personality and sensory profiles [49, 50] in the recall of spatial memory. Another study could use the two stimuli together (visual and auditory) in a third condition. A third study could focus on the suitability of the AR app for different purposes and different groups. For example, our app could be used for spatial memory training in patients with dementia and for the rehabilitation of patients with acquired brain injury.

Acknowledgements We would like to thank all of the people who participated in the study. We would like to thank the editor and reviewers for their valuable suggestions. This work was funded mainly by MCIN/AEI/10.13039/501100011033/ and by "ERDF A way of making Europe" through the project AR3Senses (TIN2017-87044-R); other support was received from Gobierno de Aragón (research group S31_20D) and FEDER 2020-2022 "Construyendo Europa desde Aragón".

Appendix: Hardware and software

For the development and validation of the app, we used a Samsung Galaxy S9 Plus with an Android 10.0 operating system. However, it can work on any device that supports ARCore[4]. The phone has two rear cameras: a main sensor of 12 MP and a secondary one for the calculation of depth of 12 MP.

To program the app, the Unity version 2018.2.1 was used as a multiplatform video game engine. The plugins of the ARCore version 1.11.0 and Azure Spatial Anchors version 1.1.1 were also incorporated into Unity. The version of ARCore used included the following features: plane detection, pose, face tracking, and light estimation. Since Azure Spatial Anchors allows augmented reality anchors to be persistently stored, our app could remember the anchors that had been placed by the supervisor when the user had to walk around the environment in the learning and evaluation phases.

ARCore and Azure Spatial Anchors work together to recognize the environment and to place the stimuli. ARCore recognizes the planes to place the objects, and Azure Spatial Anchors scans the environment looking for characteristic points (not necessarily planes) and creates a cloud of points that is stored in Azure. ARCore and Azure Spatial Anchors are related. The scanning of the environment begins when the first object is placed. A spatial anchor is created at the position of this first object, and scanning is performed around that point. It is not necessary to scan the entire environment before placing the objects. This can be done at the same time as the objects are placed.

Firebase was used to store the data of the users and the placements of the objects. Firebase is a platform that is located in the cloud and integrated with Google Cloud Platform[5]. Firebase is dedicated to the development of web applications and mobile apps. Among other functions, it allows the user's access to the application to be easily managed through e-mail and other platforms such as Google, Twitter, or Facebook. In addition, Firebase offers a real-time, back-end NoSQL database, which can be used to store the app data in JSON format. This way, the data is synchronized in real time with the app and is available for consultation even when the app has no connection.

Blender version 2.8 was used for modeling or modifying the 3D objects. The models of the dog and the frog were downloaded from free3d.com. The model of the car, the bell, the phone, the violin, and the mug were downloaded from sketchfab.com. The bell, the phone, the violin, and the mug were modified using Blender. The musical note was downloaded from CGTrader. A material from Unity was applied to give the note a blue metallic appearance.

References

1. Baddeley A (1992) Working memory. Science (80-) 255:556–559. https://doi.org/10.1126/science.1736359
2. Torsten Schmidt T, Blankenburg F (2018) Brain regions that retain the spatial layout of tactile stimuli during working memory—a 'tactospatial sketchpad'? Neuroimage 178:531–539
3. Burgess N, Becker S, King JA, O'Keefe J (2001) Memory for events and their spatial context: models and experiments. Philos Trans R Soc B Biol Sci 356:1493–1503. https://doi.org/10.1098/rstb.2001.0948
4. Neguţ A, Matu SA, Sava FA, David D (2016) Task difficulty of virtual reality-based assessment tools compared to classical paper-and-pencil or computerized measures: a meta-analytic approach. Comput Human Behav 54:414–424. https://doi.org/10.1016/j.chb.2015.08.029
5. Doniger GM, Beeri MS, Bahar-Fuchs A et al (2018) Virtual reality-based cognitive-motor training for middle-aged adults at high Alzheimer's disease risk: a randomized controlled trial. Alzheimer's Dement Transl Res Clin Interv 4:118–129. https://doi.org/10.1016/j.trci.2018.02.005
6. van der Kuil MNA, Visser-Meily JMA, Evers AWM, van der Ham IJM (2018) A usability study of a serious game in cognitive rehabilitation: A compensatory navigation training in acquired brain injury patients. Front Psychol 9:846. https://doi.org/10.3389/fpsyg.2018.00846
7. Barrett AM, Muzaffar T (2014) Spatial cognitive rehabilitation and motor recovery after stroke. Curr Opin Neurol 27:653–658. https://doi.org/10.1097/WCO.0000000000000148
8. van der Ham IJM, Claessen MHG (2020) How age relates to spatial navigation performance: functional and methodological considerations. Ageing Res Rev 58:101020. https://doi.org/10.1016/j.arr.2020.101020
9. Cullen KE, Taube JS (2017) Our sense of direction: progress, controversies and challenges. Nat Neurosci 20:1465–1473. https://doi.org/10.1038/nn.4658
10. Ruddle RA, Lessels S (2009) The benefits of using a walking interface to navigate virtual environments. ACM Trans Comput Interact 16:5. https://doi.org/10.1145/1502800.1502805
11. Bigelow J, Poremba A (2014) Achilles' Ear? Inferior human short-term and recognition memory in the auditory modality. PLoS One 9:e89914. https://doi.org/10.1371/journal.pone.0089914
12. Gloede ME, Gregg MK (2019) The fidelity of visual and auditory memory. Psychon Bull Rev 26:1325–1332. https://doi.org/10.3758/s13423-019-01597-7
13. Langlois J, Bellemare C, Toulouse J, Wells GA (2015) Spatial abilities and technical skills performance in health care: a systematic

---

[4] https://developers.google.com/ar/discover/supported-devices

[5] https://edu.google.com/intl/es-419/products/google-cloud/?modal_active=none

review. Med Educ 49:1065–1085. https://doi.org/10.1111/medu.12786

14. Mitolo M, Gardini S, Caffarra P et al (2015) Relationship between spatial ability, visuospatial working memory and self-assessed spatial orientation ability: a study in older adults. Cogn Process 16:165–176. https://doi.org/10.1007/s10339-015-0647-3
15. Juan M-C, Mendez-Lopez M, Perez-Hernandez E, Albiol-Perez S (2014) Augmented reality for the assessment of children's spatial memory in real settings. PLoS One 9:e113751. https://doi.org/10.1371/journal.pone.0113751
16. Picucci L, Caffò AO, Bosco A (2011) Besides navigation accuracy: gender differences in strategy selection and level of spatial confidence. J Environ Psychol 31:430–438. https://doi.org/10.1016/j.jenvp.2011.01.005
17. Walkowiak S, Foulsham T, Eardley AF (2015) Individual differences and personality correlates of navigational performance in the virtual route learning task. Comput Human Behav 45:402–410. https://doi.org/10.1016/j.chb.2014.12.041
18. Loachamín M, Juan M-C, Mendez-Lopez M et al (2019) Developing and evaluating a game for the assessment of spatial memory using auditory stimuli. IEEE Lat Am Trans 13:1653–1661. https://doi.org/10.1109/TLA.2019.8986443
19. Bohil CJ, Alicea B, Biocca FA (2011) Virtual reality in neuroscience research and therapy. Nat Rev Neurosci 12:752–762. https://doi.org/10.1038/nrn3122
20. Fabroyir H, Teng WC (2018) Navigation in virtual environments using head-mounted displays: allocentric vs. egocentric behaviors. Comput Human Behav 80:331–343. https://doi.org/10.1016/j.chb.2017.11.033
21. León I, Tascón L, Cimadevilla JM (2016) Age and gender-related differences in a spatial memory task in humans. Behav Brain Res 306:8–12. https://doi.org/10.1016/j.bbr.2016.03.008
22. Münzer S, Zadeh MV (2016) 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 64:191–205. https://doi.org/10.1016/j.chb.2016.06.047
23. Cimadevilla JM, Lizana JR, Roldán MD et al (2014) Spatial memory alterations in children with epilepsy of genetic origin or unknown cause. Epileptic Disord 16:203–207. https://doi.org/10.1684/epd.2014.0661
24. Reggente N, Essoe JKY, Baek HY, Rissman J (2020) The method of loci in virtual reality: explicit binding of objects to spatial contexts enhances subsequent memory recall. J Cogn Enhanc 4:12–30. https://doi.org/10.1007/s41465-019-00141-8
25. Commins S, Duffin J, Chaves K et al (2019) NavWell: a simplified virtual-reality platform for spatial navigation and memory experiments. Behav Res Methods 52:1189–1207. https://doi.org/10.3758/s13428-019-01310-5
26. Cárdenas-Delgado S, Juan MC, Méndez-López M, Pérez-Hernández E (2017) Could people with stereo-deficiencies have a rich 3D experience using HMDs? In: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Pp. 97–116
27. Cárdenas-Delgado S, Méndez-López M, Juan MC, et al (2017) 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. pp. 46–57
28. Rodríguez-Andrés D, Juan M-C, Méndez-López M et al (2016) MnemoCity task: assessment of childrens spatial memory using stereoscopy and virtual environments. PLoS One 11:e0161858. https://doi.org/10.1371/journal.pone.0161858
29. Rodriguez-Andres D, Mendez-Lopez M, Juan M-C, Perez-Hernandez E (2018) A virtual object-location task for children: gender and videogame experience influence navigation; age impacts memory and completion time. Front Psychol 9:451. https://doi.org/10.3389/fpsyg.2018.00451
30. Mendez-Lopez M, Perez-Hernandez E, Juan M-C (2016) Learning in the navigational space: age differences in a short-term memory for objects task. Learn Individ Differ 50:11–22. https://doi.org/10.1016/j.lindif.2016.06.028
31. Munoz-Montoya F, Juan M-C, Mendez-Lopez M, Fidalgo C (2019) Augmented reality based on SLAM to assess spatial short-term memory. IEEE Access 7:2453–2466. https://doi.org/10.1109/ACCESS.2018.2886627
32. Munoz-Montoya F, Fidalgo C, Juan M-C, Mendez-Lopez M (2019) Memory for object location in augmented reality: the role of gender and the relationship among spatial and anxiety outcomes. Front Hum Neurosci 13:113. https://doi.org/10.3389/fnhum.2019.00113
33. Keil J, Korte A, Ratmer A et al (2020) 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 88:165–172. https://doi.org/10.1007/s41064-020-00104-1
34. Peleg-Adler R, Lanir J, Korman M (2018) The effects of aging on the use of handheld augmented reality in a route planning task. Comput Human Behav 81:52–62. https://doi.org/10.1016/j.chb.2017.12.003
35. Chu CH, Wang SL, Tseng BC (2017) Mobile navigation services with augmented reality. IEEJ Trans Electr Electron Eng 12:S95–S103. https://doi.org/10.1002/tee.22443
36. Rehman U, Cao S (2017) Augmented-reality-based indoor navigation: a comparative analysis of handheld devices versus google glass. IEEE Trans Human-Machine Syst 47:140–151. https://doi.org/10.1109/THMS.2016.2620106
37. Cattaneo Z, Bhatt E, Merabet LB et al (2008) The influence of reduced visual acuity on age-related decline in spatial working memory: an investigation. Aging Neuropsychol Cogn 15:687–702. https://doi.org/10.1080/13825580802036951
38. Papadopoulos K, Koustriava E (2011) The impact of vision in spatial coding. Res Dev Disabil 32:2084–2091. https://doi.org/10.1016/j.ridd.2011.07.041
39. Calle-Bustos A-M, Juan M-C, García-García I, Abad F (2017) An augmented reality game to support therapeutic education for children with diabetes. PLoS One 12:e0184645. https://doi.org/10.1371/journal.pone.0184645
40. Brooke J (1996) SUS-A quick and dirty usability scale. In: Jordan PW, Thomas B, Weerdmeester BA, McClelland AL (eds) Usability evaluation in industry. Taylor & Francis, London
41. Regenbrecht H, Schubert T (2002) Measuring presence in augmented reality environments: design and a first test of a questionnaire. In: Proc 5th Annu Int Workshop Presence. Pp. 1–7
42. Slater M, Usoh M, Steed A (1994) Depth of presence in virtual environments. Presence Teleoperators Virtual Environ 3:130–144. https://doi.org/10.1162/pres.1994.3.2.130
43. Witmer BG, Singer MJ (1998) Measuring presence in virtual environments: a presence questionnaire. Presence Teleoperators Virtual Environ 7:225–240. https://doi.org/10.1162/105474698565686
44. Patrício M, Ferreira F, Oliveiros B, Caramelo F (2017) Comparing the performance of normality tests with ROC analysis and confidence intervals. Commun Stat Simul Comput 46:7535–7551. https://doi.org/10.1080/03610918.2016.1241410
45. Munoz-Montoya F, Juan MC, Mendez-Lopez M et al (2021) SLAM-based augmented reality for the assessment of short-Term spatial memory. A comparative study of visual versus tactile stimuli. PLoS One 16:1–30. https://doi.org/10.1371/journal.pone.0245976
46. Davis SW, Zhuang J, Wright P, Tyler LK (2014) Age-related sensitivity to task-related modulation of language-processing networks. Neuropsychologia 63:107–115. https://doi.org/10.1016/j.neuropsychologia.2014.08.017

47. Koen JD, Borders AA, Petzold MT, Yonelinas AP (2017) Visual short-term memory for high resolution associations is impaired in patients with medial temporal lobe damage. Hippocampus 27:184–193. https://doi.org/10.1002/hipo.22682
48. Hampstead BM, Stringer AY, Stilla RF et al (2011) Where did I put that? Patients with amnestic mild cognitive impairment demonstrate widespread reductions in activity during the encoding of ecologically relevant object-location associations. Neuropsychologia 49:2349–2361. https://doi.org/10.1016/j.neuropsychologia.2011.04.008
49. Dunn W (1997) The impact of sensory processing abilities on the daily lives of young children and their families: a conceptual model. Inf Young Child 9(4):23–35
50. Metz AE, Boling D, DeVore A et al (2019) Dunn's model of sensory processing: an investigation of the axes of the four-quadrant model in healthy adults. Brain Sci 9:35. https://doi.org/10.3390/brainsci9020035

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