Cárdenas-Delgado, S.; Méndez-López, M.; Juan, M.; Pérez-Hernández, E.; Lluch, J.; Vivó, R. (2017). Using a Virtual Maze Task to assess spatial short-term memory in adults. SCITEPRESS. 46-57. https://doi.org/10.5220/0006093200460057

Using a Virtual Maze Task to Assess Spatial Short-term Memory in Adults

Sonia Cárdenas-Delgado[1], Magdalena Méndez-López[2], M.-Carmen Juan[1], Elena Pérez-Hernández[3],
Javier Lluch[1] and Roberto Vivó[1]
[1] *Instituto Universitario de Automática e Informática Industrial, Universitat Politècnica de València, Camino de Vera, s/n, 46022 València, Spain*
[2] *Departamento de Psicología y Sociología, Universidad de Zaragoza, Zaragoza, Spain*
[3] *Departamento de Psicología Evolutiva y de la Educación, Universidad Autónoma de Madrid, Madrid, Spain*
scardenas@dsic.upv.es, mmendez@unizar.es, mcarmen@dsic.upv.es, elena.perezh@uam.es, {jlluch, rvivo}@dsic.upv.es

Keywords: Short-term Memory, Spatial Memory, VR HMD, Virtual Reality, Interdisciplinary Projects.

Abstract: In this paper, we present the Virtual Maze Task that assesses spatial short-term memory in adults involving physical movement and immersion. For physical movement, we used a real bicycle. For immersion, we used a VR HMD. We compared the exposure to the task using two different interaction types (physical active vs. physical inactive conditions). The performance and sensations of the participants were compared in both conditions. We also compared the performance on the virtual task with classical neuropsychological tests. A total of 89 adults participated in our study. The participants' ability to learn a route within the Virtual Maze Task was tested. Then, the participants assessed their experience scoring the following aspects: interaction and satisfaction. The data were analyzed and we found no differences in satisfaction and interaction scores between the physical active and the physical inactive conditions. However, the condition used for interaction affected the score obtained in the task. There were also significant effects of gender and/or interaction used in other measures of performance on the task. Finally, the performance on the task correlated with the performance on other classical neuropsychological tests for the assessment of short-term memory and spatial memory.

1 INTRODUCTION

The evolution of technology has affected all fields, including psychology. This evolution includes improvements in hardware and software for the development of more immersive experiences. According to several experts, Virtual Reality (VR) has the potential to become one of the top breakthrough technologies of the next decade (The Farm 51, 2015).

Traditionally, paper and pencil tests and computerized tests have been used for the assessment of cognitive skills. However, computer-based environments, especially VR systems for neuropsychological assessment, represent a major advance for the assessment of cognitive skills in a more ecological way. Ecological validity refers to the degree to which test results relate to real-life performance (Chaytor and Schmitter-Edgecombe, 2003).

With regard to the use of VR applied to psychology, some authors highlighted the possibility of using VR measures for neuropsychological assessment in research applications as well as in clinical practice (Negut et al., 2016). The neuropsychological assessment of individual cognitive skills improves our understanding of individual differences in behavior and helps us to detect pathology (Lezak, 1995). The study of Juan et al. (2014) showed the advantages of using an Augmented Reality application for the assessment of spatial memory in children. Spatial memory is an important cognitive skill for survival because it allows us to find our way in environments and is related to a wide range of cognitive abilities. Their study focused on the assessment of short-term memory like ours, which can be defined as the capacity for holding a small amount of information in mind in an active state for a short period of time. The children who participated in the study were satisfied with the application and considered that it was easy to use. In addition, the application developed by the study, was a valid tool for assessing the spatial short-term memory ecologically (Juan et al., 2014).

In real life, the vestibular and visual systems receive stimuli from the real environment. However, in VR, vestibular information may not be present or be influenced by optical flow patterns that are characteristic of self motion (Hettinger and Riccio, 1992). In this work, we determine if physical movement (directly related to the vestibular system) has a significant influence on spatial memory.

In this paper, we present the Virtual Maze Task that assesses spatial short-term memory in healthy adults involving physical movement and immersion. For movement, we used a real bicycle. For immersion, we used a VR Head-Mounted Displays (HMD). Specifically, we used the Oculus Rift DK2. However, the task will work exactly the same with other brand of HMD. The objective of the study was to test the ability of the new task to assess spatial short-term memory by comparing the participants' performance for the developed task with current approaches for testing spatial short-term memory. Moreover, the new task includes two types of interaction (physical active vs. physical inactive). In the physical active condition, the participant rode a bicycle. In the physical inactive condition, the participant used a gamepad.

Our VR task is based on egocentric orientation. The user learns ones body position in space for orientation (i.e., idiothetic information). Kelly et al. (Kelly and Mcnamara, 2008) also studied how egocentric experience, intrinsic structure, and extrinsic structure interact in a virtual environment. They found that the acquisition of spatial knowledge is similar to using virtual and real environments. In addition, the study suggested that VR has advantages for studying spatial memory and allows for the ease of creating different environments.

The primary hypothesis of this work was that the Virtual Maze Task could evaluate short-term spatial memory and spatial orientation in adults like the traditional procedures applied in psychology. The second hypothesis is that there would be no statistically significant difference in the score of the task between genders. The third hypothesis is that there would be no statistically significant difference for the score of the task between the two types of condition. The fourth hypothesis is that there would be no statistically significant differences in the satisfaction and interaction of the task between the two types of interaction.

This paper is structured as follows. Section 2 mentions related works. Section 3 focuses on the description of the virtual environment and the software and hardware used. Section 4 presents the sample, the measures considered, and the procedure of our study in detail. Section 5 describes the results. Section 6 presents the discussion. Finally, Section 7 summarizes the study and mentions future lines of work.

2 RELATED WORK

With regard to ecological validity, Canty et al. (Canty et al., 2014) evaluated the sensitivity, convergent validity and ecological validity of a virtual reality task for assessing prospective memory (i.e., the Virtual Reality Shopping Task). The task was tested with patients who have suffered a traumatic brain injury. They developed a VR shopping center and used a laptop screen to visualize the environment. Their results showed that the task was sensitive and ecologically measured the time and events based on prospective memory ability in patients with post-traumatic brain injury. That work allowed them to prove the benefits of using VR in the assessment and rehabilitation of memory in individuals with traumatic brain injury. Plancher et al. (Plancher et al., 2012) used a laptop to present a three-dimensional view of two urban environments inspired by Paris. In addition, a soundtrack of typical city noises (cars, people, etc.) were added to give the participants the feeling of being immersed in each environment. The participants were seated on a chair, and the virtual environment was projected 150 cm in front of them. The environment was explored by means of a virtual car using a real steering wheel, a gas pedal and a brake pedal. The results demonstrated that complex virtual environments may provide tools to reflect subjective cognitive deficits in pathological aging. The study also demonstrated the feasibility of using VR technology to study the episodic memory deficits of patients with amnesic mild cognitive impairment and Alzheimer's disease.

Although most studies use conventional monitors for showing the virtual environment, Parsons and Rizzo (Parsons and Rizzo, 2008) used a HMD eMagic Z800 to assess and compare the psychometric properties between the virtual environment and paper-and-pencil measures. They created a Virtual Reality Cognitive Performance Assessment Test. Their test focused on neurocognitive testing using a virtual city to assess recall of targets delivered within the city. Their findings revealed that there were significant correlations between the total memory score of their test and the classical learning and memory tests. In this line, Nori et al. (Nori et al., 2015) developed a VR test based on the WalCT, which is a test for assessing memory for sequences of steps within a real setting (Piccardi et al., 2008). That test aimed to assess human navigational ability. They used a HMD eMargin z800, and a graphic Workstation HP. Participants had to learn 8-step sequences, which were shown by an avatar. Their results showed that there were no differences between the real version and the virtual version of the same test. They also indicated that the virtual test was a good tool for studying the brain networks involved in sequential topographical learning.

Overall, most of the works have used simple and static stimuli. In our study, we created a Virtual Maze Task for assessing spatial short-term memory in adults. The task includes two types of conditions (physical active and physical inactive) to control navigation while participants are immersed in the virtual world. The Oculus Rift DK2 was used in our task. The Oculus Rift has already been used as a visualization device for different purposes. For example, Space Rift that is a VR game, taught children about the solar system by allowing them to explore it in a virtual environment (Peña and Tobias, 2014). Space Rift was tested with fifth-grade students. The students described the game as enjoyable and immersive, although they had problems distinguishing some of the images due to lack of sharpness.

Other works have compared different versions of the same virtual environment using the Oculus Rift. For example, two different virtual roller coasters were compared, each with different levels of fidelity (Davis et al., 2015). They found that the more realistic roller coaster with higher levels of visual flow had a significantly greater chance of inducing cybersickness. Oculus Rift DK2 was used for watching movies in which two conditions were considered: the observer condition, in which the participant was observing the scene as in traditional movies; and the actor condition, in which the participant was observing from the perspective of one of the actors and he/she became part of the plot (Van den Boom et al., 2015). They only found differences between the two conditions with regard to spatial presence in favour of the actor condition.

The Oculus Rift has also been compared with different visualization systems. For example, the Oculus Rift and a high-cost Nvis SX60 HMD were compared, which differ in resolution, field of view, and inertial properties, among other factors (Young et al., 2014). They also assessed simulator sickness and presence. The findings showed that the Oculus Rift consistently outperformed the Nvis SX60 HMD, but some people were more subject to simulator sickness with the Oculus Rift. A nVisor MH60V HMD, the Oculus Rift DK1, and Samsung Gear VR were used to learn anatomy with students of medical disciplines (Buń et al., 2015). Twenty students from the Poznan University of Technology participated in a study concerning perception. The participants were asked to select the preferred HMD and interaction method. Most of them chose the Gear VR in combination with Kinect and the gamepad as the preferred solution. Tan et al. (Tan et al., 2015) presented a study involving 10 participants that played a first-person shooter game using the Oculus Rift and a traditional desktop computer-monitor. They concluded that the participants had heightened experiences, a richer engagement with passive game elements, a higher degree of flow, and a deeper immersion with the Oculus Rift than on a traditional desktop computer-monitor. However, they also mentioned the problems of cybersickness and lack of control. Gutiérrez-Maldonado et al. (Gutiérrez-Maldonado et al., 2015) developed a VR system to train diagnostic skills for eating disorders and compared two visualization systems (Oculus Rift DK1 vs. a laptop with a stereoscopic 15.6-inch screen). In this study, fifty-two undergraduate students participated. No differences were found in either effectiveness or usability with regard to skills training in psychopathological exploration of eating disorders through virtual simulations.

3 VIRTUAL MAZE TASK

The Cincinnati water maze is a commonly accepted tool for assessing the spatial memory in rodents (see Figure 1) (Arias et al., 2014). The advantages of using this maze to detect memory impairments in rodents have been noted (Vorhees and Makris, 2015). We created a virtual maze based on the Cincinnati water maze to assess spatial memory in humans (see Figure 2). The original Cincinnati water maze has nine intersections. Our maze also has nine intersections and four of these intersections were modified to increase complexity. The maze has a wall of hedges that are two meters high and pathways of grass that are two meters wide (see Figure 3). Different animals were placed on the route (e.g., a butterfly, a tortoise, a snail or a bird). These animals helped to learn the route within the maze. Each animal is placed in different positions and at different heights. All the animals are located on the right side of the route. For example, the

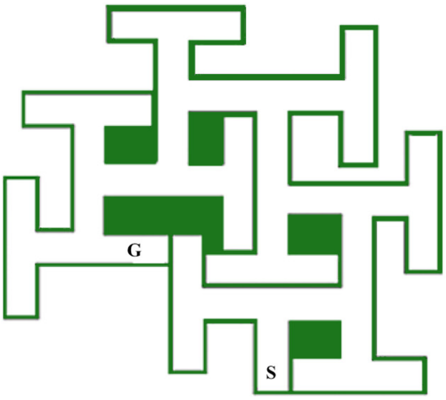

Figure 1: Schematic drawing of the Cincinnati Water Maze.

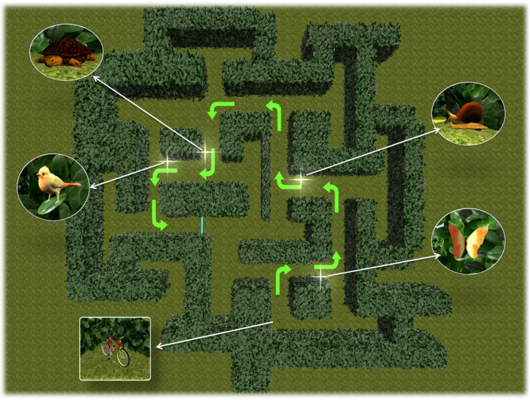

Figure 2: Virtual environment with the 3D animals. Maze viewed from above.

butterfly is placed in the second intersection and at the top of the wall; the snail is placed in the fourth intersection and on the route. The positions and heights of the animals are always the same.

We included a 3D bicycle for the navigation in the virtual maze. This bicycle is integrated in the virtual environment as an avatar from a first-person perspective. The bicycle represents the participant's point of view and personifies his/her movements in the maze. The participant is able to move around the virtual maze while he/she is observing the animals. The movements are controlled so that the participant only goes forward. The participant can turn right or left by using the handlebar of the 3D bicycle. The route is indicated by using arrows. Two types of arrows are used. The green arrows are used in the learning stage. The yellow arrows are used in the testing stage. Therefore, the green arrows are used to guide the participant from the beginning to the end of the maze. They show options at each intersection. The goal of this stage is for the participant to learn the route. In the testing stage, the yellow arrows only appear to announce that the user is near an intersection. These arrows disappear when the participant has chosen the correct path. The arrows appear again at the next intersection. If the participant chooses the wrong way, he/she is automatically placed back at the starting point, and he/she has to start the testing stage again.

For the interaction type, two different mechanisms are integrated. The first type of interaction uses a gamepad and the participant is seated on a chair (physical inactive condition). The participant is also able to move forward and to turn right or left by using the controller of the gamepad. The participant only has to stop moving the controller of the gamepad to stop the 3D bicycle. The second type of interaction uses a physical bicycle (physical active condition).

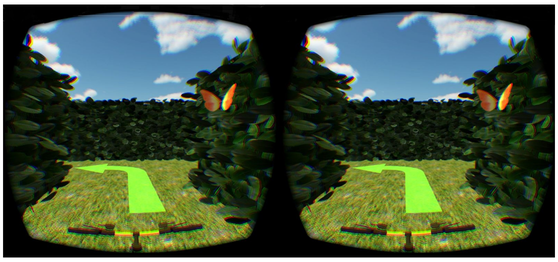

Figure 3: View of the virtual maze.

The participant controls the navigation and the 3D bicycle by using the physical bicycle. When he/she pedals the physical bicycle, he/she moves forward in the virtual maze. The participant is also able to control the turns by using the handlebar. When the participant wants to stop the bicycle, he/she only has to press the brake of the physical bicycle. All of these effects are also reproduced in the 3D bicycle.

We integrated the Oculus Rift DK2 for the immersion of the participant in our VR maze. The Oculus Rift DK2 was used in the two conditions (physical active and physical inactive). The features of this HMD greatly contribute to the level of immersion.

3.1 Hardware and Software

The Virtual Maze Task ran on an Intel Core i7 computer, 3.5 GHz processor with 16 GB RAM, an NVIDIA GeForce GTX-970 with a video card of 4GB, and Windows 8 Operating System. For the development of the virtual environment, we used Unity Edition Professional (http://unity3d.com), version 4.6.0f3, as the game engine. Blender, version 2.72, was used to create the 3D models of the animals that were included in the environment.

The Virtual Maze Task has five scenes that were created with Unity and programmed with C# and Javascript. The first scene is for the introduction of the participant's data. The person in charge of the exposure introduces the participant's date of birth and gender and chooses the type of interaction. Then, the system assigns a different code to each participant. The second scene has an option menu for choosing the stage of assessment and/or to return to the first scene. The third scene is the Habituation stage. The fourth scene is the Learning stage. The fifth scene is the Testing stage. When the participant finishes a stage, another menu is displayed allowing the participant to continue to the next stage until the task is complete. These stages are explained in Section 4.2. Two loudspeakers are used to provide messages and instructions to the participants.

The HMD for the visualization was an Oculus Rift

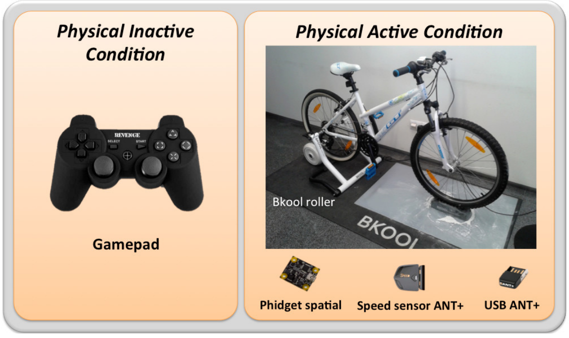

Figure 4: Mechanisms and devices of interaction.

Figure 5: A participant carrying out the task with the bicycle.

DK2. This device has a resolution of 960 × 1080 per eye, a field-of-view of 100 nominal, a weight of 0.32 kg, an optical frame rate of 75 Hz, head tracking, and positional tracking. This device also has an HDMI connector that needs to be plugged into the HDMI port of the graphics card of the computer. Video is sent to the Oculus Rift by this cable. This device also includes a USB, which carries data and power to the device, and an audio jack 2.5 mm located on the side. In addition, this version includes an external IR camera that tracks the position of the participant's head in the 3D space. A detailed description can be found in (Desai et al., 2014).

To integrate the Oculus Rift with the Virtual Maze Task, we used the plugins provided by the manufacturer (Oculus SDK 0.4.2, Oculus Runtime, and Oculus Unity Integration Package). Once the *OculusUnityIntegration.unitypackage* has been imported into Unity, the following components are available: *OVRMainMenu*, *OVRPlayerController*, and *OVRCameraController*. The Unity *MainCamera* has to be replaced by *OVRCamera*. The scripts required for the integration of the Oculus Rift DK2 with the virtual environment were programmed with C#.

For the interaction type, two different mechanisms are used. The first mechanism is a Gamepad model AB-Move BG Revenge (see Figure 4). We used the functions of the Controller Input Manager of Unity in order to integrate the gamepad with the Virtual Maze Task.

For the second mechanism of interaction, a GT mountain bike was used (see Figures 4 and 5). We use a Bkool roller (http://www.bkool.com) with an ANT+ Bike Cadence Sensor. In fact, Bkool is a smart bicycle trainer. The Bkool trainer features an advanced cycling simulator and a powerful analysis platform. In our case, we only use Bkool roller (classical model) to obtain the cadence, and to fix the rear wheel of the bicycle (see Figure 5). The model of Bkool used includes a base for the front tire and a black rubber mat (1820 × 810 × 6 mm). The rubber mat has a gel core that provides the elasticity needed to protect the floor and ensure the stability of the roller and of the bike. A fiber plate was placed on the rubber mat so that the base can rotate easily. In addition, we used a speed sensor for the bike (MyCiclo Speed Sensor ANT+$^{\text{TM}}$). This speed sensor was attached to the chain stay using plastic ties. It uses ANT+$^{\text{TM}}$ wireless transmission technology. The data received by the speed sensor is transmitted to the antenna ANT+$^{\text{TM}}$ wireless receiver (USB ANT+$^{\text{TM}}$) that connects to the computer. The sensations are very real. It allows real feelings of pedaling and offers high stability. The noise emitted is low (about 76 dB).

A program with Visual C++ was developed to manage the speed data received. This program detects which device is connected to synchronize data transmission between the receiving antenna and the ANT+$^{\text{TM}}$ sensor, and receives speed data.

The accelerometer board 1056 PhidgetSpatial 3/3/3 from Phidgets (http://www.phidgets.com) was used to obtain the rotation of the handlebar of the bicycle. It was connected to the computer via a USB cable. The dimensions are 36 × 31 × 6 mm. The libraries (*Phidget21 Libraries Setup*) supplied by the manufacturer were installed to set this accelerometer. A program with Visual C++ was developed in order to obtain the angles of rotation of the handlebar. This program also checks that the device is connected, sends error messages, initializes the accelerometer, updates data, and obtains the position data in 3D and the angle in radians.

A script was developed in C#, called Interaction, to control the movement of the avatar from the data obtained using the two mechanisms of interaction. In the case of interaction using the bicycle, the script uses information from the speed and rotation of the handlebar to calculate the rotation angle of the handlebar "n" degrees relative to the y-axis. It uses the *Quaternion.Euler* method. The *Transform.Rotate* method is used to activate the rotation of the virtual

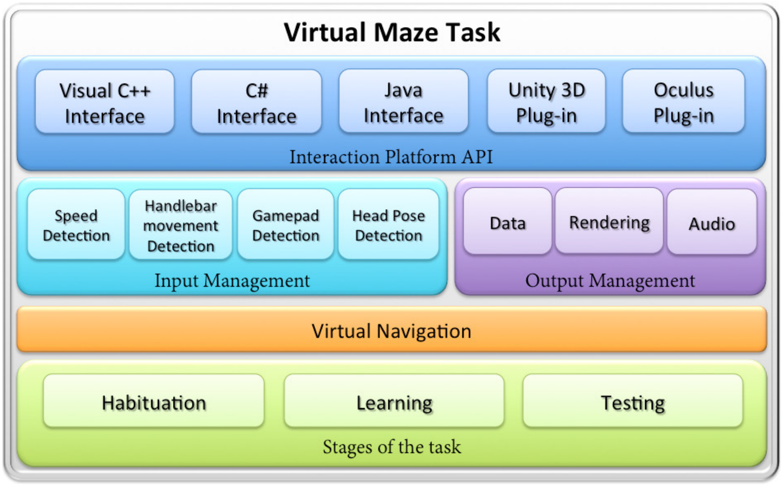

Figure 6: Scheme of the Virtual Maze Task.

bicycle (avatar). The bicycle speed data are used by the *Controller.Move* method to activate the forward movement and displacement of the virtual bicycle in the environment.

The *Input.GetAxis* method (“axisName”) was used in the Interaction. The *Input.GetAxis* returns values in the range [-1,+1]. The neutral position is 0. When axisName = “Horizontal” returns]0,+1] for the horizontal movement to the right and]0,-1] for the horizontal movement to the left. For the vertical axis, the axisName = “Vertical” and the functionality is similar. A condition that disables the backward movement was added so that the user cannot move in reverse. The values of the horizontal movement of the gamepad lever are used by the *Quaternion.Euler* method to calculate the rotation of the handlebar. The *Transform.Rotate* method is used to activate the rotation of the virtual bike (avatar). The values of the vertical movement of the gamepad lever are used by the *Controller.Move* method to activate the forward movement and the displacement of the virtual bicycle in the environment. A general scheme of the Virtual Maze Task can be seen in Figure 6.

4 STUDY

All of the participants were informed in writing about the aims and procedures of the study, and they signed an informed consent form. They were fully free to leave the study at any time, and the study was conducted according to the principles stated in the Declaration of Helsinki. The Ethics Committee of the Universitat Politècnica de València (Spain) approved the research protocol.

4.1 Participants

University students participated in this study (N=92). A recruitment campaign was conducted to find the participants by advertising within the campus facilities. The participants were randomly assigned to one of the following conditions: physical active condition (N=47) and physical inactive condition (N=45). Three participants did not finish the task due to that they presented symptoms of cybersickness (in physical active condition were 2 women and physical inactive condition was 1 men). These three participants were excluded from the sample. Therefore, the total of participants considered for our study was 89: physical active condition (N=45) and physical inactive condition (N=44). The 89 participants completed the task and filled out the questionnaires. The mean age in the physical active condition was 26.38 ± 3.87 years old and the mean age in the physical inactive condition was 25.38 ± 4.11 years old. There were 25 women and 20 men in the physical active condition, and 21 women and 23 men in the physical inactive condition. We determined the handedness of the participants (Oldfield, 1971). In the physical active condition, 39 participants were right-handed, 1 participant was left-handed and 5 participants were ambidextrous. In the physical inactive condition, 34 participants were right-handed, 7 participants were left-handed, and 3 participants were ambidextrous. The participants filled out a questionnaire, which provided information about habits with the aim of controlling variables that could interfere in the interpretation of the results. The participants did not have habits (drugs and medications) that could influence our study. Also, they did not have symptoms of sickness before the task, based on the Simulator Sickness Questionnaire (SSQ) (Kennedy and Lane, 1993).

The education levels of the participants in the physical active condition in percentages were the following: undergraduate students (42.9%), graduate students (23.8%), Master's students (21.4%), and PhD students (9.5%). In the physical inactive condition, the education levels of the participants were: undergraduate students (27.9%), graduate students (23.2%), Master's students (34.9%), and PhD students (7.0%).

4.2 Measures and Procedure

In our study, spatial short-term memory was assessed by testing the participants' ability to learn a route within a maze. The virtual maze described in the Section 3, was used for this purpose.

The Virtual Maze Task has three stages: habituation, learning, and testing (see Figure 6). The habituation stage has an environment with a short route for approximately one minute. The path has four intersections and a straight road at the end. This is a trial stage to train participants to handle the system properly and to check that the Oculus Rift DK2 is properly positioned on their heads. The learning stage consists of an environment in which the participant follows another route with nine intersections and is guided by green arrows. The participant must learn the path. The testing stage has yellow arrows that show options at each intersection. The participant must remember and follow the same route that was followed in the learning stage. The participants were immersed in the virtual maze as if they were riding a bicycle. They could see the landscape and identify the animals to determine their positions. When the participants make a mistake in the choice of the direction, the system shows a warning message and they are automatically relocated back to the starting position. Each participant has five attempts to reach the end of the maze. The time increases with the number of attempts. The experience lasts around six minutes. The average time for the interaction with the bike was 6.52 minutes, and the average time for the interaction with the gamepad was 5.27 minutes. However, the time could increase based on the number of attempts.

As measures of performance on the Virtual Maze Task, we calculated the following: the number of attempts to successfully complete the path in the testing stage (VMAttempts), the time for completion of the testing stage in seconds (VMTime), the number of participant's head turns performed at intersections in which he/she chose a correct direction during the testing stage (VMHeading), and the score (VMScore). The VMScore was obtained by adding the number of correct directions chosen in each of the five attempts established to complete the path in the testing stage. We defined ten points per attempt and a maximum VMScore of fifty points.

Spatial ability was also assessed with classical neuropsychological tests. We administered the Corsi Blocks Task (CBT, forward (CBTF) and backward (CBTB) versions), which assessed visuospatial short-term working memory (Kessels et al., 2000). We also assessed verbal short-term working memory. For this purpose, we used two verbal span subtests of the TOMAL battery: Digits Forward (DF) and Digits Backward (DB) (Reynolds and Bigler, 1994). The DF is a number recall task that measures low-level rote recall of a sequence of numbers. The DB task (a variation of the DF task) consists of a recall of a sequence of numbers, but in reverse order. For the assessment of left-right orientation ability, we used a paper pencil adaptation of the computerized Random Walker Test (RWT) (Uchiyama et al., 2009). We used the verbal version of the RWT, which provides verbal instructions, and the participants must judge the spatially correct direction. The score and the time for completion were used as measures of performance on the RWT, we used the acronym RWTS and RWTT, respectively. We considered the direct scores for the CBTF, CBTB, the DF and the DB subtests.

The participants were tested individually in two sessions, which took place on the same day. The participants were randomly assigned to one of the following experimental sessions: Session I and Session II. In Session I, the participants were assessed with the Virtual Maze Task, and they then were evaluated with neuropsychological tests. In Session II, the participants were evaluated with neuropsychological tests, and they then were assessed with the Virtual Maze Task. The different steps of the experimental procedure are shown in Figure 7.

Before starting the testing sessions, each participant was verbally informed about the exposure session, the virtual environment, the Oculus Rift DK2 as

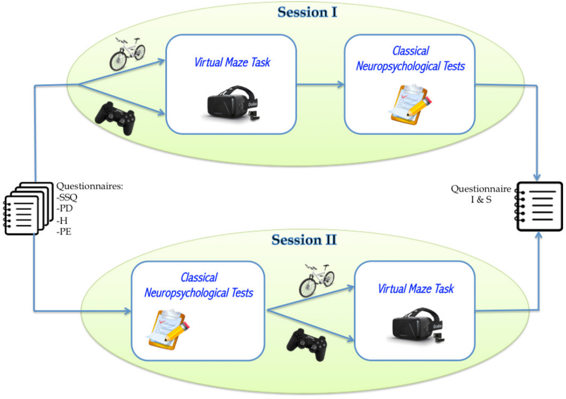

Figure 7: Procedure of the Virtual Maze Task.

the visualization device, and the type of interaction used. Also, each participant completed the handedness questionnaire, the questionnaire about personal data, and the SSQ. When he/she finished the virtual task, he/she answered another questionnaire on the interaction and satisfaction (Table 1). We also assessed the previous experience of the participants with 3D technology and other technological devices (Table 2).

Table 1: Questionnaire on interaction and satisfaction with the environment. The questionnaire used a Likert scale [from 1 to 5 (1 being *'none'* or extremely low and 5 being *'very high'*)].

#QI	**Interaction**
QI1	The environment was easy to use.
QI2	How natural was the mechanism that controlled movement through the environment?
QI3	How responsive was the environment to actions that you initiated (or performed)?
QI4	How natural did your interactions with the 3D environment seem?
QI5	How closely were you able to examine objects?
QI6	How well could you examine objects from multiple viewpoints?
QI7	In general, rate the experience of movement and interaction with the virtual environment.
#QS	**Satisfaction**
QS1	Would you use this environment another time?
QS2	How much fun did you have?
QS3	Would you invite your friends to use the environment?
QS4	Score the game from 1 to 5.
QS5	My 3D experience compared to other previous 3D experiences has been ...

Table 2: Questionnaire on previous experiences with 3D technology and other technological devices. The questionnaire used a Likert scale [from 1 to 5 (1 being *'none'* or extremely low and 5 being *'very high'*)].

#QPE	**Previous experiences**
QPE1	I play video games on computer, mobile phone, ...
QPE2	I perform activities in 3D.
QPE3	I play 3D games.

5 RESULTS

This section presents the analysis of the data collected from this study. The statistical program SPSS, version 20 (SPSS Inc., USA, 2011) was used to conduct all statistical analyses. In order to explore means and standard deviation, an initial descriptive analysis was carried out. First, data normality was checked. Our data fit the normal distribution. Therefore, the tests used were parametric. A one-way analysis of variance (ANOVA) was performed to analyze questionnaire responses regarding interaction and satisfaction outcomes. A two-way ANOVA was conducted which examined the effect of gender and interaction on the Virtual Maze Task results. Pearson's correlations were carried out to explore the relationship between Virtual Maze Task measures and neuropsychological tests. For all of the tests, a $p < .05$ determined significance.

5.1 Interaction and Satisfaction Outcomes

The responses to each question about interaction (QI) were averaged to yield a composite score for interaction (7 items, $\alpha = .693$). We did the same for the questions about satisfaction (QS) (5 questions, $\alpha = .789$) and the questions about previous experiences with 3D technology and other technological devices (QPE) (3 items, $\alpha = .530$). As Table 3 shows, no statistically significant differences were found for any of the interaction and satisfaction questions. Similarly, there were no differences between the two groups considering previous experiences.

Table 3: Mean ± Standard Deviation for the composite score about interaction (QI), satisfaction (QS), and previous experiences (QPE). One-way ANOVA between the physical active condition (Bike) and the physical inactive condition (Gamepad) and *r* effect size.

	Bike	**Gamepad**	***(F)*** ***p-value***	***r***
QI1-QI7	3.87 ± 0.49	4.08 ± 0.52	(3.85) .053	0.042
QS1-QS5	4.05 ± 0.61	4.12 ± 0.69	(0.28) .595	0.003
QPE1-QPE3	1.74 ± 0.67	1.98 ± 0.72	(2.75) .101	0.031

5.2 Virtual Maze Task Outcomes and Correlations with Neuropsychological Tests

A two-way ANOVA (Gender × Interaction) was used to analyze the measures obtained in the Virtual Maze Task. The results are shown in Table 4. Men performed a higher number of attempts to complete the testing stage than women. Also, the participants who used the bike made more attempts than those who used the gamepad. There were no differences between men and women or between conditions assigned for the time spent to complete the testing

Table 4: Mean ± Standard Deviation for measures obtained in the Virtual Maze Task by men and women in the physical active condition (Bike) and physical inactive condition (Gamepad). Two-way ANOVA (Gender × Condition). The asterisk (*) indicates significant differences.

	Condition				**Effect**		
	Bike		**GamePad**		**Gender**	**Condition**	**Interaction**
Measures	**Men**	**Women**	**Men**	**Women**	***(F)*** ***p-value***	***(F)*** ***p-value***	***(F)*** ***p-value***
VMAttempts	2.25 ± 1.74	0.80 ± 1.22	0.78 ± 1.24	0.95 ± 0.74	(5.55) .02*	(5.85) .02*	(8.88) .004*
VMTime	382 ± 136	328 ± 179	290 ± 167	379 ± 130	(0.26) .61*	(0.39) .53	(4.66) .034*
VMHeading	2.45 ± 1.93	1.08 ± 0.91	0.87 ± 0.92	1.38 ± 1.11	(2.57) .11	(5.72) .02*	(12.37) .001*
VMScore	44.75 ± 5.68	46.64 ± 4.34	48.26 ± 2.86	47.00 ± 3.25	(0.13) .72	(4.82) .03*	(3.19) .080

stage. The men who used the physical active condition made more head turns. Finally, the participants who used the physical inactive condition scored better than those who performed in the physical active condition.

The results of the correlations found between the Virtual Maze Task measures and the performance scores on classical neuropsychological tests are shown in Table 5. There are significant correlations between our task and classical tests.

6 DISCUSSION

In our work, the capability of our Virtual Maze Task was tested to assess spatial short-term memory in adults. Some applications for assessing spatial memory in humans have been developed previously (Koening et al., 2011; Cánovas et al., 2011). These applications used basic methods of human computer interaction. A review of the literature indicates that a task that incorporates stereoscopy (VR HMD) and physical movement (ride a bike) for the assessment of spatial short-term memory has not yet been developed.

The significant correlations found between the performance on our virtual task and classical neuropsychological tests suggest that our task involved sustained attentional demands and higher working memory capacity. These results also corroborate our primary hypothesis. Based on the correlation with the RWT, egocentric orientation also played a significant role in the performance of this VR task (Uchiyama et al., 2009). The positive relation with the DF and DB could suggest that verbal strategies contributed to solving the task, helping to verbally memorize the body turns associated with choice points and the landmarks (Spiers and Maguire, 2008). The negative correlation found between the head turns made at intersections and the score on the task was interesting. This result reinforces the possibility of the verbal strategy being a better strategy than other types, such as memorizing the body turns. In line with this, it should be pointed out that the Oculus Rift was a good tool for the assessment of the position of the participant's head in the 3D space, providing us with valuable information that has not been considered in other studies with virtual mazes (Werkhoven et al., 2014; Zancada-Menendez et al., 2015). This information helps us to understand the factors that contribute to learning in complex spatial environments.

Differences in the Virtual Maze Task score were not statistically significant for gender. This result corroborates our second hypothesis. However, the Virtual Maze Task score showed statistically significant differences between the two types of conditions, in favour of the physical inactive condition. This result does not corroborate the third hypothesis. We expected that there would be no differences and that if they had been, they were in favor of the physical active condition. As mentioned in the introduction section, the physical movement is directly related to the vestibular system and we hypothesized that it would have a positive influence on spatial memory. However, this influence has not been reflected in the results. Although unexpected, this result in favour of the physical inactive condition is in line with the study of Cutmore et al. (Cutmore et al., 2000), which found that spatial learning in virtual environments with an active exposure was not more advantageous than a passive exposure. Also, the differences for type of interaction show the importance of methodological factors in the study of spatial memory in humans (Andreano and Cahill, 2009). Moreover, our physical inactive condition can also be performed by people with reduced mobility (Hill-Briggs et al., 2007).

The participants did not differ in their opinions

Table 5: The correlation matrix of the Virtual Maze Task and classical neuropsychological test performance scores. The correlation coefficients (*r*) that reached significance (*p: p-value*) are displayed in bold type.

		VMTime	**VMHeading**	**VMScore**	**CBTF**	**CBTB**	**DF**	**DB**	**RWTS**	**RWTT**
VMAttempts	*r*	**.59**	**.62**	**-.48**	**-.21**	**-.25**	-.17	**-.24**	**-.29**	.17
	p	**<.0001**	**<.0001**	**<.0001**	**.04**	**.02**	.11	**.02**	**.005**	.11
VMTime	*r*		**.59**	**-.33**	-.19	-.16	**-.26**	**-.21**	-.17	.19
	p		**<.0001**	**.002**	.07	.13	**.01**	**.04**	.11	.07
VMHeading	*r*			**-.51**	.01	-.02	-.16	-.19	-.12	-.01
	p			**<.0001**	.92	.87	.14	.09	.28	.92
VMScore	*r*				.18	**.28**	**.40**	**.38**	**.31**	**-.30**
	p				.08	**.008**	**<.0001**	**<.0001**	**.003**	**.003**
CBTF	*r*					**.73**	**.36**	**.23**	.10	**-.53**
	p					**<.0001**	**<.0001**	**.03**	.36	**<.0001**
CBTB	*r*						**.35**	**.31**	**.26**	**-.60**
	p						**.001**	**.003**	**.02**	**<.0001**
DF	*r*							**.61**	.09	**-.41**
	p							**<.0001**	.39	**<.0001**
DB	*r*								.16	**-.33**
	p								.12	**.001**
RWTS	*r*									-.17
	p									.11

about interaction and satisfaction with the experience in the Virtual Maze Task. These results corroborate our fourth hypothesis.

From our point of view, the current HMDs (e.g., Oculus Rift) have many possibilities. As mentioned in the introduction section, the Oculus Rift has already been used in psychology. For example, Gutiérrez-Maldonado et al. (Gutiérrez-Maldonado et al., 2015) used it for training diagnostic skills in eating disorders. Based on the results obtained, we believe that the Oculus Rift and other HMDs have great potential for psychology, especially for the assessment of spatial short-term memory.

Even though the Oculus Rift has several benefits, it also has some drawbacks. One of the drawbacks to our proposal is that the Oculus Rift DK2 needs a computer connection by wire. The use of a wireless VR HMD with the same or greater immersion features would make a freer system that would allow the user freedom of movement without fear of stumbling upon or becoming tangled in cables. According to some predictions (The Farm 51, 2015), half a billion VR headsets will be sold per year by 2025, and more than 400 hundred million will be wireless VR HMD. In these predictions, the number of wireless VR HMD sold in 2016 is more or less the same as the wired VR HMD. However, this trend is not predicted to continue. It has been predicted that a hundred million of VR HMDs will be sold by 2020. Of these, less than 20% will be wired VR HMDs. We share this opinion and think that the wireless VR HMD would be decisive in the future for many applications. Another drawback of the Oculus Rift (in general of the HMDs) is the cybersickness that the HMDs may induce. As Davis et al. (Davis et al., 2015) indicated, the more realistic the environment with higher levels of visual flow, the greater the chance of inducing cybersickness. It would be very interesting to determine whether the Oculus Rift induces more cybersickness than other HMDs. Cybersickness is a limitation in our task. In fact, 3 out of 92 participants in our study did not finish the task. Therefore, people prone to cybersickness could not use this type of tasks. Another limitation of our physical active condition is for people with mobility problems.

Initially, we used the Wii Remote controller to obtain the turns of the handlebar of the bicycle. However, since the Oculus Rift DK2 is used for visualization, there was a conflict between the two devices that made their simultaneous use impossible due to that both use infrared sensor. This must be taken into account in future developments.

7 CONCLUSIONS

We have developed a new Virtual Maze Task to assess spatial short-term memory in adults. We compared the performance of the new task with traditional neuropsychological procedures, and we measured the usability and satisfaction of the participants for the new task. The performance in the Virtual Maze Task was compared to other tests of spatial and memory skills. According to a measure of overall execution, the performance on the new task was better in the participants who used in the physical inactive condition than in the physical active condition. However, the usability and satisfaction did not differ between conditions. These results showed that the type of interaction used is a relevant methodological issue in studies about cognition that are based on VR technologies. The Virtual Maze Task could be used as an entertaining method to assess or train adults in spatial short-term memory skills.

The Cincinnati water maze has commonly been used in studies with rodents. In our task, the Cincinnati water maze has been visualized using the Oculus Rift and tested with human adults. Our study and other previous works (e.g., Cánovas et al., 2011) support the potential of VR for adapting tasks developed for animals to humans.

For future work, a study to compare HMDs of others models and brands, taking into account their features such as resolution, field-of-view, and latency can be considered. We would also like to study the capability of the Virtual Maze Task to detect learning difficulties in samples of people with academic problems or neurological disorders. The possibilities of our task for children could also be studied. In fact, we are currently testing a different virtual environment using a large stereo screen with polarization glasses for the assessment of spatial memory in children. However, other devices could also be used, paying special attention to wireless HMDs, such as Samsung Gear VR. Other types of interaction could be studied (e.g., gesture interaction).

ACKNOWLEDGEMENTS

- This work was mainly funded by the Spanish Ministry of Economy and Competitiveness (MINECO) through the CHILDMNEMOS project (TIN2012-37381-C02-01) and co-financed by the European Regional Development Fund (FEDER).
- Other financial support was received from the Government of Aragon (Department of Industry and Innovation), the European Social Fund for Aragon, and the Government of the Republic of Ecuador through the Scholarship Program of the Secretary of Higher Education, Science, Technology and Innovation (SENESCYT).
- We would like to thank the following for their contributions:
  - BKOOL for lending us the Bkool roller.
  - This work would not have been possible without their collaboration: Mauricio Loachamín Valencia, David Rodríguez Andrés, and Juan Fernando Martín.
  - DSIC and ASIC for allowing us to use its facilities during the testing phase, especially to Vicente Blasco and Manuel Jiménez.
  - The users who participated in the study.
  - The reviewers for their valuable comments.

REFERENCES

Andreano, J. M. and Cahill, L. (2009). Sex influences on the neurobiology of learning and memory. *Learning and Memory*, 16:248–266.

Arias, N., Méndez, M., and Arias, J. L. (2014). Brain networks underlying navigation in the Cincinnati water maze with external and internal cues. *Neuroscience Letters*, 576:68–72.

Buń, P., Górski, F., Wichniarek, R., Kuczko, W., Hamrol, A., and Zawadzki, P. (2015). Application of professional and low-cost head mounted devices in immersive educational application. In *Procedia Computer Science*, pages 173–181.

Cánovas, R., García, R. F., and Cimadevilla, J. M. (2011). Effect of reference frames and number of cues available on the spatial orientation of males and females in a virtual memory task. *Behavioural Brain Research*, 216(1):116–121.

Canty, A. L., Fleming, J., Patterson, F., Green, H. J., Man, D., and Shum, D. H. (2014). Evaluation of a virtual reality prospective memory task for use with individuals with severe traumatic brain injury. *Neuropsychological Rehabilitation*, 24(2):238–265.

Chaytor, N. and Schmitter-Edgecombe, M. (2003). The ecological validity of neuropsychological tests: A review of the literature on everyday cognitive skills. *Neuropsychology Review*, 13:181–197.

Cutmore, T. R. H., Hine, T. J., Maberly, K. J., Langford, N. M., and Hawgood, G. (2000). Cognitive and gender factors influencing navigation in a virtual environment. *International Journal of Human-Computer Studies*, 53(2):223–249.

Davis, S., Nesbitt, K., and Nalivaiko, E. (2015). Comparing the onset of cybersickness using the Oculus Rift and two virtual roller coasters. *11th Australasian Conference on Interactive Entertainment (IE 2015)*, pages 27–30.

Desai, P. R., Desai, P. N., Ajmera, K. D., and Mehta, K. (2014). A review paper on Oculus Rift-A virtual reality headset. *International Journal of Engineering Trends and Technology*, 13(4):175–179.

Gutiérrez-Maldonado, J., Ferrer-García, M., Pla-Sanjuanelo, J., Andrés-Pueyo, A., and Talarn-Caparrós, A. (2015). Virtual Reality to train diagnostic skills in eating disorders. Comparison of two low cost systems. *Studies in Health Technology and Informatics*, 219:75–81.

Hettinger, L. J. and Riccio, G. E. (1992). Visually induced motion sickness in virtual environments. *Presence: Teleoperators & Virtual Environments*, 1(3):306–310.

Hill-Briggs, F., Dial, J. G., Morere, D. A., and Joyce, A. (2007). Neuropsychological assessment of persons with physical disability, visual impairment or blindness, and hearing impairment or deafness. *Archives of Clinical Neuropsychology*, 22(3):389–404.

Juan, M. C., Mendez-Lopez, M., Perez-Hernandez, E., and Albiol-Perez, S. (2014). Augmented reality for the assessment of children's spatial memory in seal settings. *PLoS ONE*, 9(12):e113751.

Kelly, J. W. and Mcnamara, T. P. (2008). Spatial memories of virtual environments: How egocentric experience, intrinsic structure, and extrinsic structure interact. *Psychonomic Bulletin & Review*, 15(2):322–327.

Kennedy, R. and Lane, N. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *The International Journal of Aviation Psychology*, 3(3):203–220.

Kessels, R. P. C., van Zandvoort, M. J. E., Postma, A., Kappelle, L. J., and de Haan, E. H. F. (2000). The Corsi Block-Tapping Task: Standardization and normative data. *Applied Neuropsychology*, 7(4):252–258.

Koening, S., Crucian, G., Dünser, A., Bartneck, C., and Dalrymple-Alford, J. (2011). Validity evaluation of a spatial memory task in virtual environments. *International Journal of Design and Innovation Research*, 6:1–13.

Lezak, M. D. (1995). *Neuropsychological assessment*. 3rd ed., Oxford University Press, New York, NY.

Negut, A., Matu, S. A., Sava, F. A., and David, D. (2016). Virtual reality measures in neuropsychological assessment: A meta-analytic review. *Clinical Neuropsychology*, 30(2):165–184.

Nori, R., Piccardi, L., Migliori, M., Guidazzoli, A., Frasca, F., De Luca, D., and Giusberti, F. (2015). The virtual reality walking Corsi test. *Computers in Human Behavior*, 48:72–77.

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, 9(1):97–113.

Parsons, T. D. and Rizzo, A. A. (2008). Initial validation of a virtual environment for assessment of memory functioning: Virtual reality cognitive performance assessment test. *CyberPsychology & Behavior*, 11(1):17–25.

Peña, J. G. V. and Tobias, G. P. A. R. (2014). Space Rift: An Oculus Rift solar system exploration game. *Philippine IT Journal*, 7(1):55–60.

Piccardi, L., Iaria, G., Ricci, M., Bianchini, F., Zompanti, L., and Guariglia, C. (2008). Walking in the Corsi test: Which type of memory do you need?. *Neuroscience Letters*, 432(2):127–131.

Plancher, G., Tirard, A., Gyselinck, V., Nicolas, S., and Piolino, P. (2012). Using virtual reality to characterize episodic memory profiles in amnestic mild cognitive impairment and Alzheimer's disease: Influence of active and passive encoding. *Neuropsychologia*, 50(5):592–602.

Reynolds, C. R. and Bigler, E. D. (1994). *TOMAL Test of memory and learning: Examiner's manual*. Austin, TX Pro-Ed [In TOMAL Test de memoria y aprendizaje. Manual de interpretación (E. Goikoetxea & Departamento I+D de TEA Ediciones, Adapters), 2001, Madrid, Spain: TEA Ediciones].

Spiers, H. J. and Maguire, E. A. (2008). The dynamic nature of cognition during wayfinding. *Journal of Environmental Psychology*, 28:232–249.

Tan, C. T., Leong, T. W., Shen, S., Dubravs, C., and Si, C. (2015). Exploring gameplay experiences on the Oculus Rift. In *Proceedings of CHI Play '15*, pages 253–263.

The Farm 51 (2015). *Report on the current state of the VR market*. In The Farm 51. Group S.A., http://thefarm51.com/ripress/VR_market_report_2015_The_Farm51.pdf. Accessed 2016 December 06.

Uchiyama, H., Mitsuishi, K., and Ohno, H. (2009). Random Walker Test: A computerized alternative to the Road-Map Test. *Behavior Research Methods*, 41(4):1242–53.

Van den Boom, A. A., Stupar-Rutenfrans, S., Bastiaens, O. S. P., and Van Gisbergen, M. M. S. (2015). *Observe or participate: The effect of point-of-view on presence and enjoyment in 360 degree movies for head mounted displays*. http://ceur-ws.org/Vol-1528/paper13.pdf. Accessed 2016 December 06.

Vorhees, C. V. and Makris, S. L. (2015). Assessment of learning, memory, and attention in developmental neurotoxicity regulatory studies: Synthesis, commentary, and recommendations. *Neurotoxicology and Teratology*, 52:109–115.

Werkhoven, P., van Erp, J. B. F., and Philippi, T. G. (2014). Navigating virtual mazes: The benefits of audiovisual landmarks. *Displays*, 35(3):110–117.

Young, M. K., Gaylor, G. B., Andrus, S. M., and Bodenheimer, B. (2014). A comparison of two cost-differentiated virtual reality systems for perception and action tasks. In *Proceedings of the ACM Symposium on Applied Perception*, pages 83–90.

Zancada-Menendez, C., Sampedro-Piquero, P., Meneghetti, C., Labate, E., Begega, A., and Lopez, L. (2015). Age differences in path learning: The role of interference in updating spatial information. *Learning and Individual Differences*, 38:83–89.