BRIEF REPORT

NavWell: A simplified virtual-reality platform for spatial navigation and memory experiments

Sean Commins [1] (ORCID) · Joseph Duffin [2] · Keylor Chaves [2] · Diarmuid Leahy [2] · Kevin Corcoran [2] · Michelle Caffrey [1] · Lisa Keenan [1] · Deirdre Finan [2] · Conor Thornberry [1]

Abstract

Being able to navigate, recall important locations, and find the way home are critical skills, essential for survival for both humans and animals. These skills can be examined in the laboratory using the Morris water maze, often considered the gold standard test of animal navigation. In this task, animals are required to locate and recall the location of an escape platform hidden in a pool filled with water. Because animals can not see the platform directly, they must use various landmarks in the environment to escape. With recent advances in technology and virtual reality (VR), many tasks originally used in the animal literature can now be translated for human studies. The virtual water maze task is no exception. However, a number of issues are associated with these mazes, including cost, lack of flexibility, and lack of standardization in terms of experimental designs and procedures. Here we present a virtual water maze system (NavWell) that is readily downloadable and free to use. The system allows for the easy design of experiments and the testing of participants on a desktop computer or fully immersive VR environment. The data from four independent experiments are presented in order to validate the software. From these experiments, a set of procedures for use with a number of well-known memory tests is suggested. This potentially can help with the standardization of navigational research and with navigational testing in the clinic or in an educational environment. Finally, we discuss the limitations of the software and plans for its development and future use.

Keywords Navigation · Memory · Spatial · Virtual reality · Water maze

Knowing how to get from a to b, recalling important places and recognizing one's current location are critical skills for everyday living. How animals and humans navigate has been an intense research topic across many decades. Such studies have led to theoretical debates on the nature of spatial memories, including cognitive mapping theory (O'Keefe & Nadel, 1979) or associative learning accounts (Pearce, 2009), as well as, elucidating the role played by various brains structures

(including the hippocampus, entorhinal, parietal, and retrospenial cortices; Morris, Garrud, Rawlins, & O'Keefe, 1982) in successful navigation. The discovery of place, grid, head-direction and other specialized cells (Grieves & Jeffery, 2017) has also provided a unique insight into the neural mechanism of navigation. In addition to these findings, impairment in spatial memory and navigation is a feature of many disorders and normal ageing decline (Head & Isom, 2010). For example, people living with schizophrenia (Kargar, Askari, Khoshaman, & Mohammadi, 2019), stroke (Takahashi, Kawamura, Shiota, Kasahata, & Hirayama, 1997), acquired brain injury (Kolarik et al., 2016), and Alzheimer's disease (Zanco et al., 2018) all show spatial deficits. Furthermore, as spatial disorientation is a key deficit with mild cognitive impairment and a strong predictor of AD, there is increasing interest in the possibility of using navigation as an early marker for this disease (Coughlan, Laczó, Hort, & Hornberger, 2018). Indeed it has only recently been recognized that navigation assessment has tremendous potential to be used not only as a predictor of future AD, but as an outcome measure

✉ Sean Commins
  Sean.Commins@nuim.ie

1 Department of Psychology, Maynooth University, County Kildare, Ireland

2 Department of Computer Science, Maynooth University, County Kildare, Ireland

in behavioral and pharmacological intervention studies, as well as, a possible cognitive-training task itself (see Lester, Moffat, Wiener, Barnes, & Wolbers, 2017, for details).

The gold standard in animal spatial memory testing is the Morris water maze (Morris, 1981). In this task animals are required to find a platform, hidden just below the surface in a large circular pool of water. As animals can not see the goal directly they must use various cues in the environment to locate the target and recall its location on subsequent trials. The task offers advantage over other mazes as it eliminates the use of odor cues and the water provides a strong motivator for escape. The time taken to find the hidden platform is used as a measure of spatial navigation and location memory. Other measures including distance traveled, time spent in various regions of the pool and heading may also be used. The task also offers great flexibility; reference memory, spatial working memory, procedural memory, and cued learning can all be examined using this task by simply altering the protocol used. The water maze task relies on an intact hippocampus (Diviney, Fey, & Commins, 2013; Morris et al., 1982) and is therefore a good behavioral measure for pharmacological and genetic models of diseases that are known to impact on this structure (such as schizophrenia and Alzheimer's disease; e.g., Baeta-Corral & Giménez-Llort, 2015). Furthermore, as the hippocampus is a key structure in learning and memory (Barry, Coogan, & Commins, 2016), synaptic plasticity (Bliss & Lomo, 1973; Craig & Commins, 2007), as well as, neurogenesis (Keith, Priester, Ferguson, Salling, & Hancock, 2008), changes in performance in the water maze task (or other hippocampal-dependent tasks) may reflect cellular changes that occur during normal learning processes.

Virtual reality environments are excellent for testing humans in a controlled safe environment and for translating animal research (Astur, Taylor, Mamelak, Philpott, & Sutherland, 2002; Hartley, Maguire, Spiers, & Burgess, 2003). The virtual water maze task (VWM) has been successfully modeled on the animal version and has been used as a means to separate various learning strategies, to understand underlying neural substrates, to examine sex and age differences, as well as, being a sensitive assay for detecting deficits in a range of clinical populations including those with Alzheimer's disease, traumatic brain injury, and schizophrenia (Folley, Astur, Jagannathan, Calhoun, & Pearlson, 2010; Hamilton, Driscoll, & Sutherland, 2002; Newhouse, Newhouse, & Astur, 2007; Possin et al., 2016; Skelton, Ross, Nerad, & Livingstone, 2006).

Despite the task's versatility and usefulness, a number of issues have arisen that prevents the VWM from being widely accepted as a navigational tool, these include the following. (i) Developing a VR environment from scratch is time-consuming and requires the necessary in-house expertise to which many researchers/clinicians do not have access. Furthermore, although there are a number of commercially-available VWMs these are often quite expensive. (ii) Many VR mazes are custom-designed with a single environment and set procedure, as such they are inflexible for researchers and various research interests. (iii) Related to this, many VR tasks do not offer the possibility of use across different platforms including integration with 3-D headsets. (iv) A major difficulty with many VMW tasks is that they are not standard across research groups in terms of either environments or procedures used. For example, environments of different sizes, shapes and cues numbers are used. Trial numbers and times can be radically different across research groups. As such, it is difficult to directly compare results across different groups, making generalization and replication of experiments hard to achieve. If the VMW task is to become a standard tool in research and/or the clinic it should be readily available, easy to use and come with a set of standardized procedures.

To overcome some of these issues we have developed a spatial navigation tool, NavWell, based on the water maze task that is freely downloadable onto a PC or Mac computer. With researchers in mind, a "Research" version of NavWell has been developed with an intuitive step-by-step guide allowing experimenters to create and design their own environment and testing procedures, thereby offering researchers more control and flexibility over their own experiments. Furthermore, the software is available for use with a 3-D VR headset, allowing participants to be fully immersed in the environment. As the program can be used with an android phone this allows experiments to be designed anywhere, as well as, opening the possibility of using the smartphone as a VR headset.

With clinicians and educators in mind, we have included an "Education" version of NavWell. With this version, a set of predesigned environments and protocols are included. There is no need to pick a particular environment or decide on the number of appropriate trials or cues. These have been predetermined as a result of our validation experiments (see below). A test using the visible platform, that serves to exclude motivation and motor issues, and a reference memory task have both been included in this version. This version of NavWell is particularly suited to educators and/or clinicians who may simply want to run the task as a demonstration or a test, rather than as a full experiment. Compared to the "Research" version, it is simpler to use and quicker to implement. This version is also free and easy to download, and comes with a video guide through our dedicated website.

Finally, we validated the software and procedures using a number of independent experiments; from these, we produced a number of recommended protocols for a

number of commonly used memory paradigms including reference/place memory, spatial working memory, as well as, testing of general perceptual and motor functioning. Such standardization should allow the replication and comparison of experiments across research groups, as well as, allowing a standard protocol to be used for testing in the clinic.

Workflow

NavWell procedure

The NavWell toolbox can be downloaded free from https:// navwell.cs.nuim.ie/home. From the website a "Research" or "Educational" version can be chosen. Instructions on how to download the different versions onto a Mac or PC is provided on the webpage. A full description of designing and running an experiment across various platforms in the "Research" version can be found below and by using the online tutorial (video and a pdf manual). Figure 1 shows an overview of NavWell, from designing experiments to the analysis of results.

Designing the environment

The researcher initially clicks on the NavWell_Administrator Console. Under the Environments tab, the experimenter selects "create environment." Figure 2 shows how the experimenter can create a particular environment in a step-by-step, easy-to-follow process. Steps 1 and 2 involve selection of the shape (circular, triangle, square, and rectangle) and the size (large, medium, or small) of the testing arena (Fig. 2a). Step 3 involves selecting the location of the hidden platform (small gray square in Fig. 2b; 15% of the arena), which can be placed anywhere within the arena. Step 4 involves the selection the environmental cues (from a choice of lights or geometric shapes); these can be placed on the wall anywhere in the environment (X in Fig. 2c). Once selected, the cues and their locations are displayed, but they can also be removed if the experimenter is not satisfied. This environment is then saved and can be cloned or removed, if desired.

Designing the experiment and testing participants

Once the environment has been created, the experimenter can design the experiment by selecting "create experiment" under the Experiment tag, again in a step-by-step, easy-to-follow process. Steps 1 and 2 require selection of the environment that the experimenter has just created (or other previously created environments) and selecting a starting position for each trial (red star Fig. 3a). The experimenter can repeat this for whatever number of trials

may be required. The duration of each trial and the inter-trial interval can be also chosen for each (Fig. 3b). Some experiments may require a probe trial, which is a trial in which no platform is present, usually given after learning to test retention; to achieve this, simply click the Retention trial box after specifying the trial's time. Alternatively, some experiments may require the platform to be visible; in this case, after specifying the trial length, the experimenter may simply click the Visible Trial box. A visible platform has been used in animal experiments to ensure that animals can see the environment, and this is also very useful for working with vulnerable human populations, such as those with Alzheimer disease or older participants whose vision may be compromised.

Once the environment has been created and the exact experiment designed, the experimenter can start testing particular participants. The experimenter creates a unique identification for each participant and can add names or notes as required. This is done by clicking on "add participant" under the Participant tag (Fig. 3c; experimenters should be aware of any legislative and ethical issues concerning personal data). The identification code will be used to search for and select this participant in the future. Once saved, this participant can be assigned to a particular experiment. To do this, the experimenter types the unique identity code into the search box and selects "view." This will allow the experimenter to see whether the participant has already taken part in an experiment. If so, all experiments and their results will be listed. If not, the experimenter can assign a particular experiment (or a number of experiments) from the list to this participant.

Once the experiment has been designed, the experimenter closes the NavWell_Administrator Console and opens NavWell [3/2D], which has been installed on the standalone machine or Android device. The experimenter then enters the participant ID, selects the experiment to be run, and follows the instructions on the screen. At this point, the experimenter may wish to prepare the participant with the VR headset (in the 3-D version; see below). The participant can navigate through the environment using a joystick, the arrow keys, or the "w," "a," "s," "f" keys. Individual results are saved and can be viewed. Similar to the water maze task, the time taken to reach the platform, distance traveled, and percentage of time spent in the goal quadrant are given for each trial, as well as the overall averages. In addition, visualization of individual tracks, heat maps, and the percentage of time spent in each area for every trial can be observed, noted, and saved (Figs. 1 and 4). The results for a particular experiment can be viewed under the Experiment tag. All participants that took part in a particular experiment can be viewed and saved by selecting the appropriate experiment and clicking the "Results" button. Participants' results may
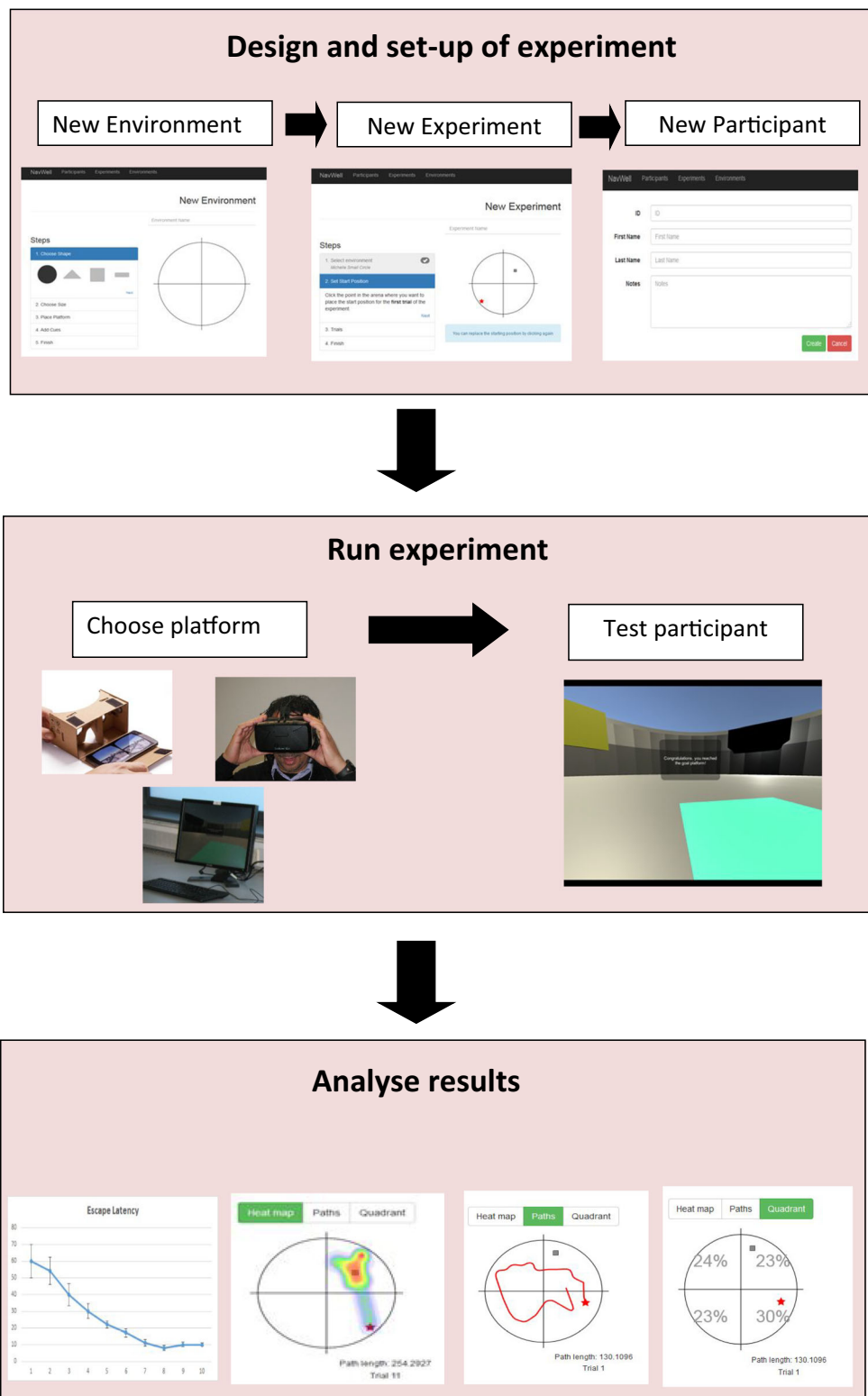
Fig. 1 Overview of NavWell, from the initial designing of environment and experiments to the testing of participants across various platforms and the subsequent analysis of results

be deleted, if appropriate. A summary of the results is then displayed, and this can be exported as a .csv file. The results from each participant within the experiment

can also be viewed by selecting "Details" beside the participant's name (see Fig. 4 for an example of the data obtained).
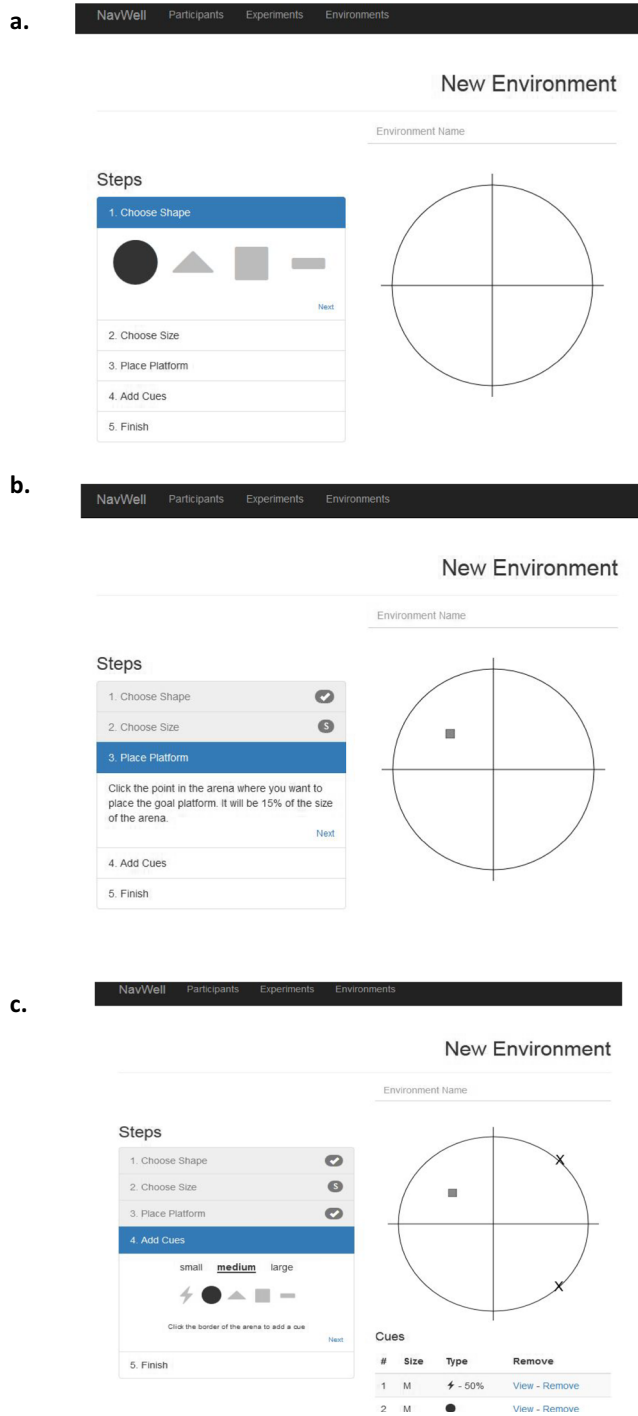
Fig. 2 Screen shots taken from NavWell, depicting the step-by-step process of designing a particular maze. Experimenters start by selecting a particular shape of the maze (a) and location of the hidden platform (b), and then they can select the various environmental cues (c) that will surround the maze
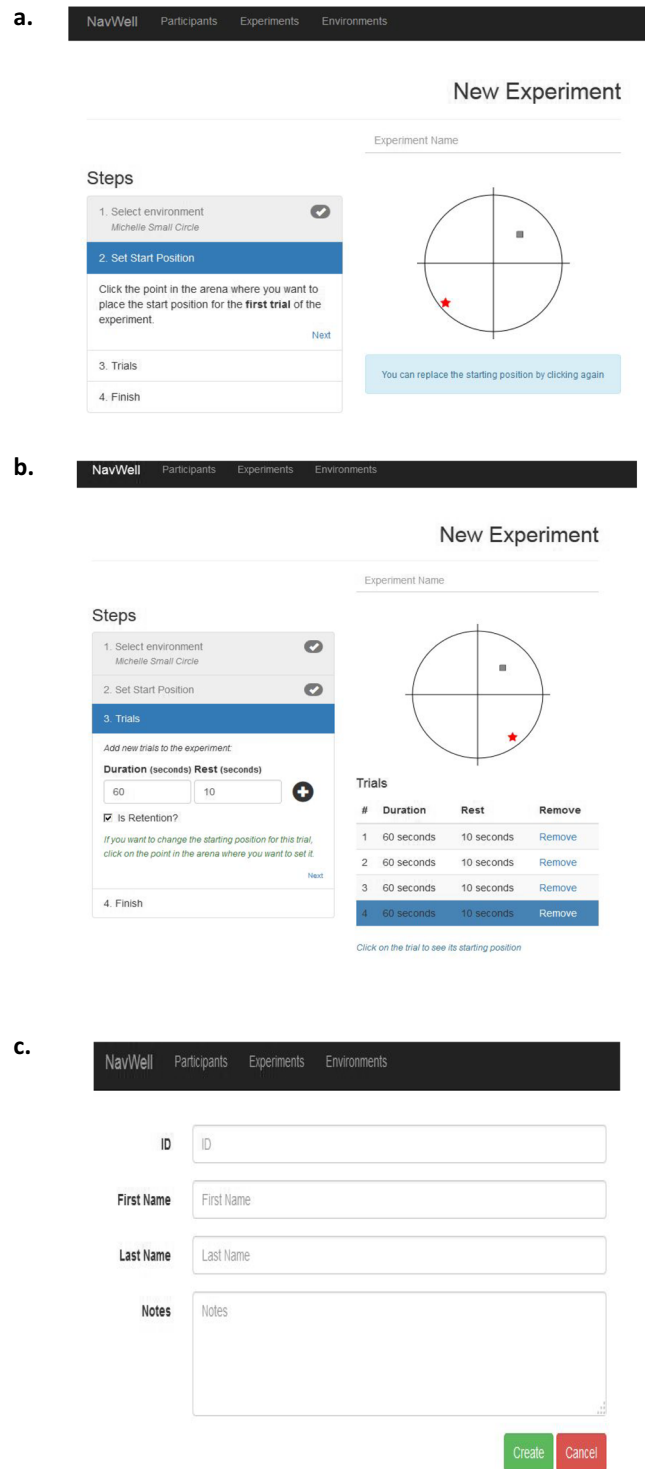

Fig. 3 Screen shots taken from NavWell, depicting the step-by-step process of designing a new experiment. The experimenter starts by selecting the environment that has already been created, and then select the starting position (a) and the number, duration, and type of trials needed for the experiment (b). Each participant can then be created and given a unique code (c)

Validation experiments

To validate and test NavWell and its procedures, a number of independent experiments were conducted. These were done to replicate previously published findings and to examine whether similar findings would be obtained by using NavWell.
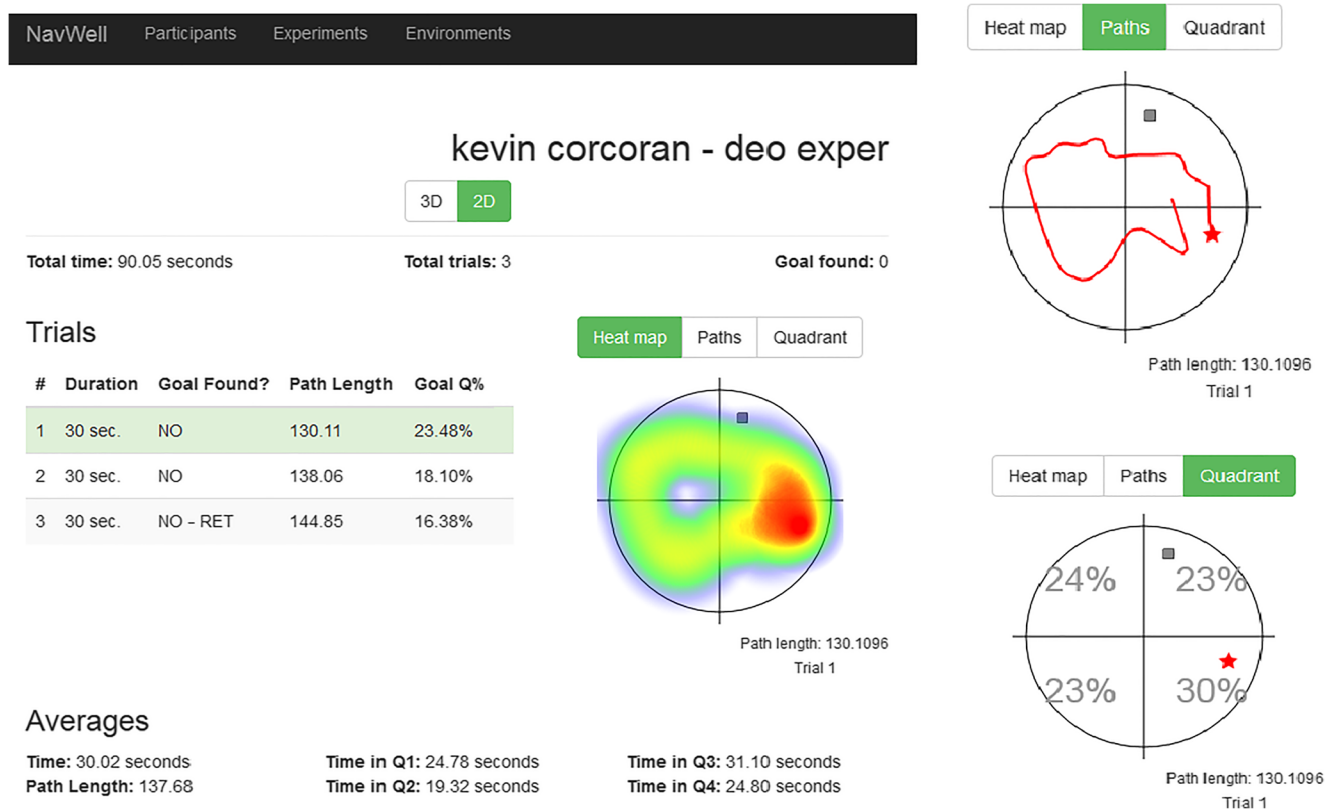
Fig. 4 Screen shots taken from NavWell, depicting the different measures that can be obtained, including the visualization of individual tracks, a heat map, and the percentage of time spent in each area for every trial and each participant

Experiment 1: Examination of spatial-learning performance between participants, using the 3-D version of the task (i.e., with VR goggles and full immersion) versus a version of NavWell displayed on a desktop computer

Brief introduction and method

Methods using 3-D environments offer a great advantage over traditional desktop displays: They are more immersive and more interactive, and they allow for greater movement. Indeed, many studies suggest that participants often have a greater sense of realism in such environments (e.g., Hupont, Gracia, Sanagustín, & Gracia, 2015). With 3-D VR becoming more common, it has been argued that more immersive environments are increasing becoming necessary for participants to accept simulations as real (Ponto, Chen, Tredinnick, & Radwin, 2014). However, there is conflicting evidence whether participants actually show better performance in more immersive tasks. For example, Witmer and Singer (1994) suggested that there is a correlation between sense of immersion and performance level, while Sousa Santos et al. (2009) showed user performance to be enhanced in nonimmersive, desktop displays. Given the importance of movement in

navigation, the possible difficulties with using technology in older adults and other vulnerable populations, we wished to compare our virtual task across both desktop and 3-D VR display.

Participants Twenty-nine participants were recruited through an ad hoc sampling of students of Maynooth University, friends and family. Each participant was randomly assigned to either the 3-D version of the task ($n = 14$) or the desktop (referred to as 2-D; $n = 15$). Both groups were matched for sex and age. The 3-D group had mean age of 21.2 ± 1.0 years (seven males and seven females). The 2-D group had a mean age of 21.7 ± 0.57 years (seven males and eight females). There was no significant difference between the groups in terms of age ($p = .529$). All participants were informed as to the nature of the experiment and were free to withdraw at any stage. Experimental protocols were approved by the Maynooth University Department of Psychology ethics committee. All participants had normal or corrected-to-normal vision. No participant experienced motion sickness using the headset or with the VR environment.

Apparatus The virtual maze in both conditions consisted of a large-sized circular environment (taking 21 s to traverse the

arena, calculated to be 29.4 virtual meters [vm] across) with the hidden platform located in the middle of the northeast quadrant. The size of the platform was 15% of the arena and consisted of a bright green square that became visible when the participant traversed it (Fig. 5a). Two cues were used and were located on the wall of the arena (both cues were lights, each at 50% luminance) in the northeast and northwest locations (see Fig. 5b). The

setup was similar to that used by Farina et al. (2015) for testing animals in the Morris water maze.

VR goggles The consumer version of the Oculus Rift version 2 (currently, NavWell does not work with HTC Vive or other devices) was used. The device consisted of a lightweight headset (470 g) that covered the entire field of view. The
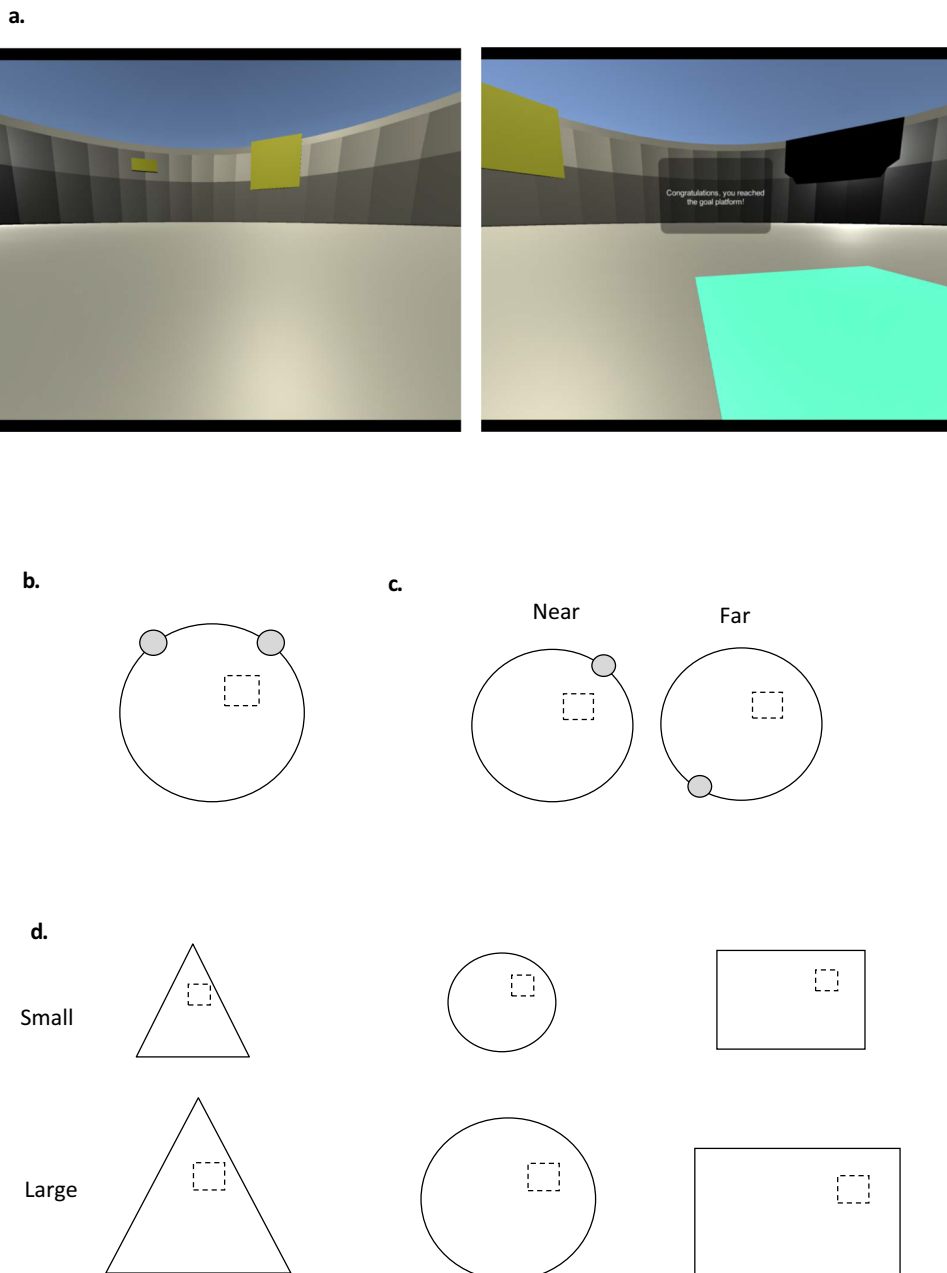

Fig. 5 (a) Screen shots taken from NavWell from the user viewpoint, as he or she navigates the environment (left) and finds the platform (right). (b) Experimental setup of the maze for the 3-D versus 2-D experiment. The dashed square represents the location of the hidden platform; the gray circles show the locations of two light cues on the wall of the arena. (c)

Experimental setup of the maze for the near/far experiment, with dashed squares indicating the location of platform, and the gray circles representing the location of the cues. (d) Experimental setup of the maze for the large/small experiment. The dashed square indicates the location of the platform in each arena of different shapes and sizes

headset had a separate display for each eye, with 1,080 × 1,200 resolution per eye and a refresh rate of 90 Hz. This headset gives a 110° horizontal field of view. Participants placed the headset on their head and used the adjustable straps so that it was comfortably secured in place with comfort. All participants used the Oculus Rift environment while sitting on a swivel chair so that they were free to rotate and look around.

Control tasks All participants were given a number of control tasks, to ensure that both groups were matched in terms of general memory ability, general intelligence, and visual–spatial processing. The first task used was the Rey auditory verbal learning task (RAVLT; Rey, 1941). This is designed to assess immediate memory span and new learning. The test consisted of 15 nouns (List A) read out loud for five consecutive trials. After each trial, the participant was asked to recall as many words as possible. The total scores from Trials 1–5 were calculated and used. The National Adult Reading Test (NART; Nelson, 1982) was also given to each participant. The NART is a widely used vocabulary-based measure of premorbid intellectual function. Participants were required to read a list of 50 phonetically irregular English words from a sheet of paper. Responses were recorded as being correct or incorrect. The number of errors (out of 50) was recorded and was used to estimate verbal IQ, performance IQ, and full scale IQ according to a conversion table. The final control measure used was the trail-making task (TMT; Reitan, 1958). This is a test used to assess task switching, sequencing, and mental flexibility, as well as visual–spatial, motor, and executive function. The test comprised Parts A and B. Part A required a participant to "join the dots" linking numbers in increasing numerical sequence going from 1 to 25. In Part B, the participant was required to "join the dots" in alternating numerical and alphabetical order starting with number "1" linked to letter "A" then linked to number "2" then to letter "B" etc. The time taken, in seconds, for each part was measured for all participants, and the difference between the parts (i.e., Trial B minus Trial A) was used.

Procedure Using NavWell, all participants were required to find the hidden platform located in the center of the northeast quadrant of the arena (Figs. 5b and 6a insets). Each participant was given 60 s to find the target. Once it was found, the participants were given 10 s to rest and then were instructed to look around the environment and note the various cues. Participants had to complete ten trials in total. For each trial, the participant started in one of four random locations (north, south, east, or west position). After this acquisition phase, the participants were given the three pen-and-paper control tasks (see above). This allowed for a time delay between acquisition

and recall. Following completion of the tasks, the participants were asked to perform a single 60-s probe/recall trial in the virtual maze, in which no platform was present. The starting position in the probe trial was from the southeast position (a novel location, different to the starting positions during training). Participants could move around the environment using a mouse pad or the arrow keys on the keyboard. The time taken to find the hidden platform and the distanced traveled were used as the dependent measures during acquisition. The percentage of time (of 60 s) was used as the dependent measure during the probe trial.

Statistics and analysis Mixed factorial analyses of variance (ANOVAs) were used to examine acquisition with Bonferroni-corrected t tests. Mixed factorial ANOVAs were used to measure the probe trials, with appropriate t tests to examine interaction effects. Significance levels were taken as $p \leq .05$.

Results

Initially, a number of independent t tests were used to compare both groups across the various control tasks. No significant difference between the two groups was found in terms of scores on the NART, with the 2-D group scoring a mean of 17.3 ± 0.6 errors and the 3-D group 17 ± 1 errors ($p = .88$). Likewise, both groups were well matched on the TMT(Trial B − A) (2-D, 21.1 ± 3 s; 3-D, 16.4 ± 2.3 s, $p = .723$) and the RAVLT (2-D, 54 ± 1.8 correct responses; 3-D, 52.9 ± 2.8 s, $p = .199$).

A 2 (group) × 10 (trial) mixed factorial ANOVA was then conducted to compare how quickly the groups reached the target during the acquisition phase of the virtual maze task. Figure 6a demonstrates that both groups learned the task readily. An overall significant effect for trial was found [$F(9, 243) = 13.42$, $p = .001$, effect size = 0.33], with participants reaching the target significantly faster on Trial 10 than on Trial 1 ($p < .05$, Bonferroni-corrected t tests). However, no group effect [$F(1, 27) = 3.195$, $p = .085$, effect size = 0.11] or interaction [$F(9, 243) = 0.781$, $p = .634$, effect size = 0.03] was noted. We then compared the total path lengths (distance, measured in virtual meters) taken by both groups across the acquisition phase. A 2 × 10 mixed factorial ANOVA demonstrated a significant effect for trial [$F(9, 243) = 12.15$, $p = .001$, effect size = 0.31], with participants showing shorter path lengths on Trial 10 than on the first trial ($p < .05$, Bonferroni-corrected t tests), but no overall group effect [$F(1, 27) = 0.048$, $p = .828$, effect size = 0.002, not shown] or Trial × Group interaction [$F(9, 243) = 0.838$, $p = .582$, effect size = 0.03] was found.

To examine retention, both groups were given 60 s to find the target. A 2 (group) × 4 (quadrant) mixed factorial ANOVA
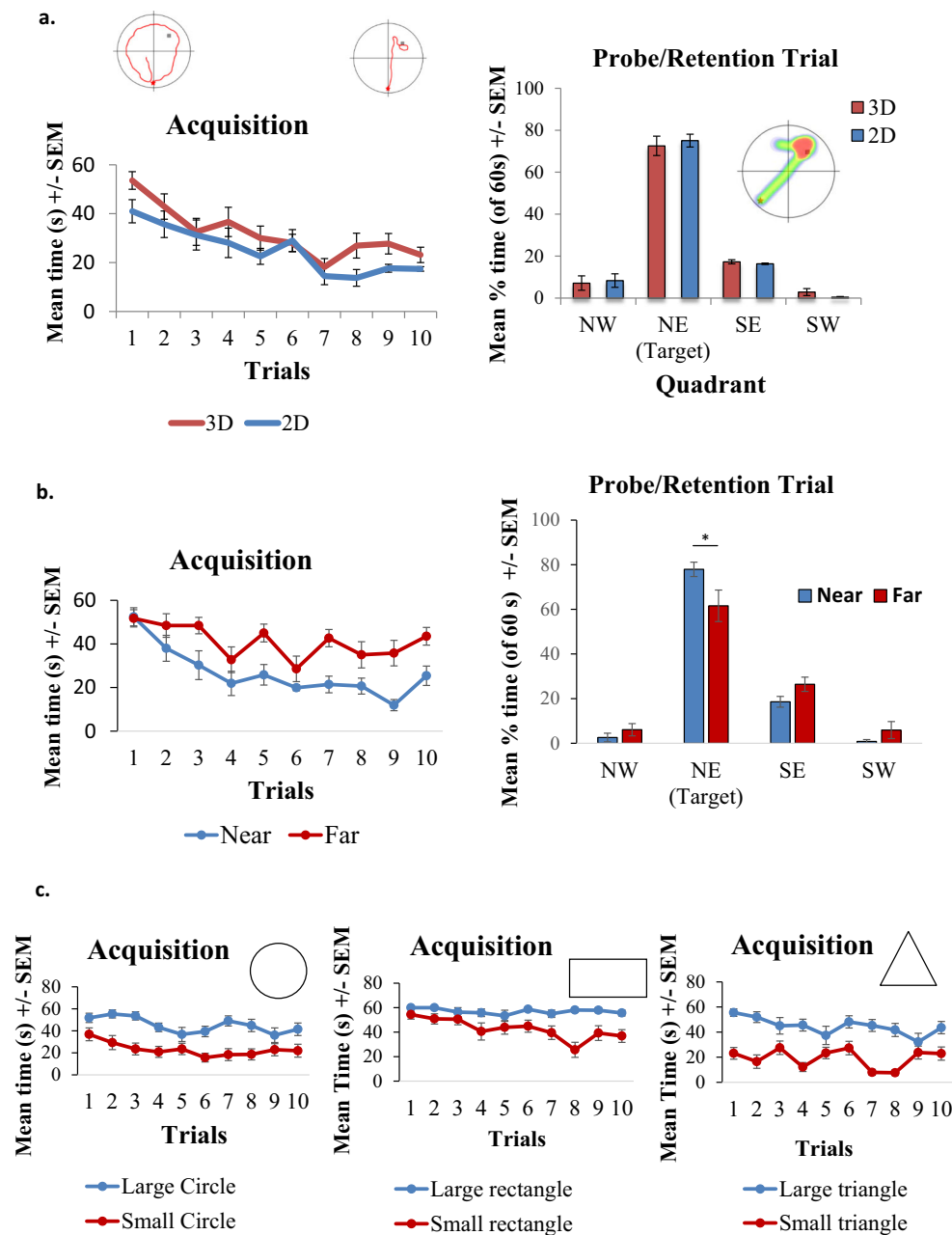
Fig. 6 (a) Comparison of latencies to the target across acquisition (left), and time spent searching for the platform during retention (right), for those using VR goggles (3-D) or using a computer (2-D) version of the virtual water maze task. (b) Comparison of acquisition (left) and retention (right) for those navigating the virtual task with a cue located in either a "far" or a "near" position (see also Fig. S4c). (c) Comparison of latencies to the target across acquisition for those in a large or a small arena of various shapes (in the absence of cues)

was conducted. We observed a significant effect of quadrant [$F(3, 81) = 236.2$, $p = .001$, effect size = 0.89], with both groups searching significantly more in the target (NE) quadrant than in the other regions ($p < .05$). A significant group effect [$F(1, 27) = 5.586$, $p = .026$, effect size = 0.17] was also noted, but there was no interaction effect [$F(3, 81) = 0.255$, $p = .857$, effect size = 0.009]. Follow-up analyses showed that that both groups searched equally in the target NE quadrant

during the probe trial, with no significant difference between the two groups [$t(27) = 0.456$, $p = .65$].

Discussion

These results suggest that NavWell can be used successfully to test spatial learning and memory, irrespective of the technology used. This finding offers researchers the flexibility to use

either a desktop or 3-D VR headset, depending on the particular research question, suitability, and availability. For example, although none of our participants reported motion sickness, which is an often-cited disadvantage of using immersive VR (e.g., Astur, Tropp, Sava, Constable, & Markus, 2004; Munafo, Diedrick, & Stoffregen, 2017). So while head-mounted displays (HMDs) may become more widespread over the coming years, many participants and patients may not be willing to use such technology if they experience dizziness, nausea and headaches. There is also some evidence that motion sickness may affect females more than males (Munafo et al., 2017), again limiting the usefulness of using such technology. The recommendation would be that any participant experiencing motion sickness should stop the task immediately; if the participant continues to feel unwell medical advice should be sought. Any health and safety issues should be considered and ethical approval should be sought before experiments commence. The use of questionnaires, for example the Simulator Sickness Questionnaire (Kennedy, Lane, Berbaum, & Lilienthal, 1993), would also be useful to assess initial level of symptoms.

Furthermore, given the expense associated with many HMDs at the current time, having a desktop version of the task that does not seem to disadvantage participants in terms of performance, provides researchers with an excellent alternative for examining spatial learning and memory. Aside from possible motion sickness, another limitation of HMDs is that it may not be suitable for all populations in a practical sense. For example, studies investigating the feasibility of immersive VRs in hospital settings have shown that older adults not only have more difficulties when using HMDs, but also tend to be less willing to participate in the first place (Mosadeghi, Reid, Martinez, Rosen, & Spiegel, 2016). There is also limited research on the effects of long-term use of VR technology with children. Other issues, such as the comfort of participants that wear glasses along with HMDs, should also be considered.

Here we demonstrated that both immersive and nonimmersive techniques are equally as effective for examining spatial navigation and memory, but care should be given when decided which version to use. For example, using HMDs may be good for research purposes but a desktop version may be more suitable in the clinic.

Experiment 2: Examination of the role of cue location on performance in spatial learning and memory using NavWell

Brief introduction and method

The use of cues or landmarks is essential for successful navigation in both the real world and virtual environments (Jansen-Osmann & Fuchs, 2006; Morris, 1981; Rodrigo, 2002). However, there is evidence that cue salience (in terms of size, brightness, and proximity to the goal) may only impact spatial performance (Commins & Fey, 2019). For example, Chamizo and Rodrigo (2004) examined the importance of relative distance of a landmark to a goal using the water maze task. They demonstrated a graded effect, whereby increasing the distance between the goal and the landmark led to a greater impairment in spatial performance. Similarly, Diviney et al. (2013) found that animals were slower to find the hidden goal when cues were placed farther away from the target (opposite the hidden goal) than when cues were located in a nearer position. This effect has also been reported using virtual environments with human participants. For example, Artigas, Aznar-Casanova, and Chamizo (2005) found that closer landmarks establish better control compared to those further away. Using the NavWell software, we wanted to replicate these animal and human findings and show that cue located further from the goal would lead to poorer spatial learning and memory. In addition, we wanted to ensure that the cues available to the participant in the NavWell environment are actually the ones that are used to navigate (as opposed to other, uncontrolled features).

Participants Twenty-six different participants took part in this experiment and were again recruited through an ad hoc sampling of students of Maynooth University, friends, and family. Each participant was randomly assigned to either a near ($n = 12$) or a far ($n = 14$) group. Similar to Experiment 1, both groups were matched for sex and age; all participants had normal or corrected-to-normal vision. The mean ages were 23.3 ± 0.9 years for the near group and 22.5 ± 0.9 years for the far group, with no significant difference between the groups ($p = .583$). Participants were informed as to the nature of the experiment and were free to withdraw at any stage.

Apparatus The virtual maze for both groups consisted of a medium circular environment (taking 15.75 s to traverse the arena, calculated at 22.05 vm), again with the platform hidden in the northeast quadrant. One cue (light of 50% luminance) was used, which was located in the northeast position for the near group and the southwest position for the far group (see Fig. 5c). The size of the platform was 15% of the arena and consisted of a bright green square that became visible when the participant crossed it.

Procedure Similar to Experiment 1, each participant was given 60 s to find the target in the NavWell task. Once found, participants were given 10 s to rest and were instructed to look around the environment and note the location of the cues. Participants were required to complete ten trials in total. For each trial the participant started in one of four random locations (north, south, east, or west position). As in Experiment 1, following the acquisition phase, participants were given pen-and-paper control tasks (see above, the NART and TMT), which took approximately 15

min. Following this, the participants were given a single 60-s probe/recall trial in the virtual maze.

Results

We first compared the two groups across the control tasks to ensure that they were matched both for general intelligence and executive function ability. No significant differences were noted between the groups for the NART (near group, 14.6 ± 1.7 errors; far group 16.3 ± 1.1 errors, $p = .393$) or the TMT(B–A) test (near group, 22 ± 3.3 s; far group 27.2 ± 4.3 s, $p = .351$).

A 2 (group) × 10 (trial) mixed factorial ANOVA was then conducted to compare how quickly the groups reached the target during the acquisition phase of the virtual maze task. An overall significant effect for trial was found [$F(9, 216) = 8.897$, $p = .001$, effect size = 0.27], demonstrating that both groups learned the task and that participants reached the target significantly faster on the final trials as compared to the first trial ($p < .05$, Bonferroni corrected t tests). A main group effect was also found, with Fig. 6b (left) demonstrating that those in the near group were significantly faster at finding the hidden platform than those in the far group [$F(1, 24) = 14.959$, $p = .001$, effect size = 0.384]. No Group × Trial interaction was found [$F(9, 216) = 1.549$, $p = .165$ effect size = 0.057]. We also compared the distance taken by both groups to reach the target across the acquisition phase (not shown). A 2 × 10 mixed factorial ANOVA demonstrated results similar to those for latency. A significant effect of trial [$F(9, 216) = 12.51$, $p = .001$, effect size = 0.343] and a significant group effect [$F(1, 24) = 8.226$, $p = .008$, effect size = 0.255] were found. No Trial × Group interaction effect was noted [$F(9, 216) = 1.930$, $p = .05$, effect size = 0.07].

Retention was assessed by allowing both groups 60 s to search for the target (without it being present). A 2 (group) × 4 (quadrant) mixed factorial ANOVA was conducted. We found a significant effect of quadrant [$F(3, 72) = 101.88$, $p = .001$, effect size = 0.81], with both groups searching significantly more in the target (NE) quadrant than in the other regions ($p < .05$). Although there was no significant group effect [$F(1, 24) = 0.219$, $p = .644$, effect size = 0.009], a Group × Quadrant interaction was shown [$F(3, 81) = 3.219$, $p = .028$, effect size = 0.118]. Follow-up analyses revealed that the near group searched in the target quadrant (NE) significantly more during the probe trial [$t(18.06) = 2.106$, $p = .05$) than did the far group (Fig. 6b right).

Discussion

This experiment demonstrated that cues are critical for successful navigation in the virtual water maze task. Furthermore, cues located further from the goal led to a poorer performance when compared to those located closer to the target. These results replicate findings from both the animal (Chamizo & Rodrigo, 2004; Diviney et al., 2013) and human (Artigas

et al., 2005) literature and helps validate NavWell as a spatial learning tool. While the far group was impaired compared to the near group, it is important to note that the far group was still able to learn the task and searched significantly more in the target quadrant during the probe trial compared to all other quadrants. There are at least two possible reasons that may help explain why participants were slower in the far group. First, the cue in the near position may have been perceived as a beacon; participants could keep the cue within their visual field, head toward it and find the goal relatively easily. Whereas, the cue in the far position may not be directly within the participants' visual field as they navigated toward the goal. As such, participants may rely more on working memory, path integration or other strategies to keep the relative position of the cue and the goal in mind, thereby increasing the cognitive load on the participant (see also Fu, Bravo, & Roskos, 2015). This increased cognitive load may translate into a slower performance. Second, Diviney et al. (2013) found in the water maze task that rats initially move toward the cues and then make a turning toward where they think the goal is located. The time taken by participants to initially head toward the far cue before readjusting may add to the slower response in this group. Although NavWell currently is able to show the exploration tracks of each individual, it does not yet have the capability of replaying the track to investigate this more idea fully. Furthermore, the software is not currently able to conduct a proper behavioral analysis on each participant's track (see, e.g., Diviney et al., 2013; Harvey, Brant, & Commins, 2009). These features will be added at a later date.

Experiment 3: Examination of the effect of arena size on spatial learning.

Brief introduction and method

Cues and landmarks are critical for successful navigation, but the environment in which these cues are set may also have an important role to play. In maze learning (virtual or otherwise), both size and shape of the environment are thought to influence performance. For example, Stankiewicz, Legge, Mansfield, and Schlicht (2006) investigated the effect of increasing layout size on spatial performance and found that participants' efficiencies decreased as the layout size increased. As well as taking longer time to navigate, a larger arena may also require a heavier cognitive and memory load, especially if participants are unable to physically see all the information available to them (Fu et al., 2015). There has also been intense research examining whether animals and humans can use the overall shape of the environment to find a hidden target. In the original experiment, Cheng (1986) trained rats to find food in a corner of a rectangular arena with each corner also marked with a distinctive feature. During retest, when the features were removed, the rats continued to search in the

correct corner as well as in the geometrically equivalent one. This suggested that animals encoded the overall shape of the environment. Similar findings have been reported in humans. Interestingly, although adults can use both feature and geometric cues, toddlers seem to use just the geometric shape and ignore distinctive features (e.g., a red wall; Hermer & Spelke, 1994, 1996). Although this experiment attempts to validate NavWell by using various shaped environments, it is not specifically looking at the question of shape and reorientation. Rather we hypothesize that, irrespective of the environment shape, participants in a larger environment should take longer to find the hidden target compared to those in a smaller environment. Although it is possible that participants may be able to use the shape of a distinctive environment (e.g., rectangle and triangle but not a circle) to aid navigation, the size effect should still hold.

Participants Twenty-six participants were recruited for this experiment from a sample of students of Maynooth University, friends, and family. Participants were randomly assigned to two different conditions: large ($n = 13$) or small arena ($n = 13$). All participants tested had normal or corrected-to-normal vision, and both groups were matched for sex and age. The mean ages for the large and small groups were 34.7 ± 4.1 years (six males/eight females) and 34.2 ± 4.6 years (six males/seven females), respectively. No significant age difference between the groups was noted ($p = .946$). All participants were informed as to the nature of the experiment, provided consent and were told that they were free to withdraw at any stage. One participant did not complete all three arenas; final numbers of participants are indicated in the sections below.

Apparatus Each participant was required to navigate and find a hidden target in a triangular, circular, and rectangular environment (either large or small, depending on the group). No cues were used in this experiment, so participants had to use the shape of the environment only. The platform was located for all conditions in the northeast region of the environment (Fig. 5d). The dimensions of the environments were as follow: small triangle, 9.28 s/12.92 vm (traversing the dissection); large triangle, 19.59 s/27.42 vm (traversing the dissection); small circle, 10.5 s/14.7 vm (traversing the diameter); large circle, 21 s/29.4 vm (traversing the diameter); small rectangle, 23.47 s/29.4 vm (traversing the diagonal); large rectangle, 42 s/58.8 vm (traversing the diagonal).

Procedure Each participant was given ten trials (60 s for each trial) to find the target, with an intertrial interval of 10 s. Participants started from a random position for each trial: north, south, east, or west for the circular environment; from each of the four corners in the rectangular environment; and from each of the three corners in the triangular environment. Following completion of the acquisition trials, each participant was given a number

of control tasks—the NART, TMT, and the RAVLT (see above). These lasted approximately 15–20 min. No retention trial was carried out in this experiment.

Results

Similar to before, we compared both the large and small group on the various control tasks using a series of independent t tests. No significant difference was found for any measure. The mean number of errors obtained on the NART for the large group was 11.5 ± 1.6, as compared to 9.6 ± 1.3 for the small group ($p = .358$). The time taken to complete the TMT(B–A) test was 22.8 ± 2.4 s for the large group and 25.6 ± 5.6 s for the small group ($p = .655$). Scores on the RAVLT were also comparable: 52.8 ± 1.6 items recalled for the large group, and 52.5 ± 1.7 items for the small group ($p = .907$).

A 2 (group) × 10 (trial) mixed factorial ANOVA was used to compare how quickly the large ($n = 13$) and small ($n = 12$) groups found the target across the acquisition phase. For the circle arena, an overall significant effect for trial [$F(9, 207) = 2.781$, $p = .004$, effect size $= 0.108$] and a large group effect [$F(1, 23) = 18.539$, $p = .001$, effect size $= 0.446$] were found. No Trial × Group interaction was noted [$F(9, 207) = 0.752$, $p = .661$, effect size $= 0.032$]. Follow-up analyses showed that although a significant effect of trial was found for the large group [$F(9, 108) = 2.075$, $p = .04$, effect size $= 0.147$], Bonferroni-corrected t tests failed to determine where that difference lay. This suggests that limited learning took place with the large circular arena. Similarly, the small group [$F(9, 99) = 1.488$, $p = .163$, effect size $= 0.119$] also failed to learn the location of the target in the circular environment (Fig. 6c).

Similar analyses compared the large ($n = 13$) and small ($n = 12$) groups in the rectangular arena. A 2 × 10 mixed factorial ANOVA demonstrated significant effects for trial [$F(9, 207) = 3.623$, $p = .001$, effect size $= 0.136$] and group [$F(1, 23) = 12.638$, $p = .002$, effect size $= 0.355$], and a Trial × Group interaction [$F(9, 207) = 2.756$, $p = .005$, effect size $= 0.107$]. Further analyses were conducted to examine each group separately. For the large rectangular group, no significant effect of trial was found [$F(9, 108) = 0.783$, $p = .632$, effect size $= 0.061$]. For the small group, a significant effect of trial was found [$F(9, 99) = 3.492$, $p = .001$, effect size $= 0.241$]. However, Bonferroni-corrected t tests showed that participants were significantly faster only on Trial 8 than on Trial 1 (Fig. 6c, middle). No other differences were noted, suggesting that limited learning took place in the rectangular environment.

Final analyses were conducted comparing participants in the large triangular arena ($n = 13$) to those in the small environment ($n = 13$). A 2 × 10 mixed factorial ANOVA demonstrated a significant effects of trial [$F(9, 216) = 3.709$, $p = .001$, effect size $= 0.134$] and group [$F(1, 24) = 28.47$, $p = .001$, effect size $= 0.543$], and a Trial × Group interaction [$F(9, 216) = 4,089$, $p = .001$, effect size $= 0.146$]. Further analyses

were conducted to examine each group separately. For the large triangular group, a significant trial effect was found [$F(9, 108) = 3.364$, $p = .001$, effect size = 0.219]; however, Bonferroni-corrected t tests failed to reveal this difference. For the small triangular group, a significant trial effect was also found [$F(9, 108) = 4.477$, $p = .001$, effect size = 0.272]; Bonferroni-corrected t tests demonstrated that some trials were significantly different from others (i.e., Trials 7 and 8 different from Trials 5 and 6; see Fig. 6c right).

Discussion

The results of this experiment demonstrated that irrespective of the environmental type (circle, rectangle, or triangle), the participants in the smaller arenas found the hidden target significantly quicker than those in the larger ones. This was to be expected and demonstrates the robustness of NavWell as a tool, since changing the environmental dimensions corresponded to a change in performance. Participants had a smaller area to navigate, and thus came across the goal much quicker and easier.

The results also demonstrate that given a simple environment, without any obvious stable landmarks, participants showed limited, if any, learning. They were unable to learn the location of the target relative to the shape of the environment. This might have been expected in the circular setup, as there is nothing in the environment to which the participants can link to the goal. Again, this is a positive finding as it shows uniformity and that the software did not produce any extra unwanted environmental factors (shadowing effects etc.) that could be used as an aid. Although, we found a significant Trial effect for some conditions, these effects were not systematic (there were differences between some trials and not others) and participants did not show convincing learning. There was no improvement in participants' performance as the number of trials increased. This finding may question the use of the overall shape of an environment to navigate accurately. For example, it has been shown that animals search for a hidden platform in two opposite locations of a rectangular arena, demonstrating their use of overall shape (Cheng, 1986). However, there are challenges to this idea. Animals may simply use knowledge about the length of one side (of a rectangle) compared to another, rather than the overall shape of the environment to help locate a target (Pearce, Good, Jones, & McGregor, 2004; see Lew et al., 2014, for similar results in humans). Angles, too, have been considered an important salient orientation cue for humans (Lubyk, Dupuis, Gutiérrez, & Spetch, 2012). Furthermore, the use of an overall geometric shape may depend on the presence and salience of local landmarks (Hayward, Good, & Pearce, 2004; Redhead & Hamilton, 2009), and it is important that such factors be ruled out. Although our results suggest that humans have difficulty learning to find a hidden target using the overall shape alone, caution must be applied,

and follow-up studies with a larger pool of participants should be done to examine the reorientation question further.

Experiment 4: Examination of the effect of age on spatial learning

Brief introduction and method

Along with many other cognitive processes, spatial memory and navigational abilities tend to decline with age. Older adults typically show poorer recall of landmarks and spatial configurations (Head & Isom, 2010) during navigation testing. For example, using a virtual water maze task, Moffat and Resnick (2002) have demonstrated that older adults travelled a greater distance and took longer to reach a goal compared to younger adults. The same authors also showed reduced activation (using fMRI) of the hippocampal region and retrosplenial cortex in older adults compared to a younger cohort, during a virtual navigation task. In contrast, older adults showed activation increases in frontal regions (Moffat, Elkins, & Resnick, 2006). Similar findings by Antonova et al. (2009) showed more bilateral hippocampal activation in younger adults compared to an older group. In addition, poorer performances by older adults on the virtual water maze task was correlated by decreased hippocampal activity. Observed age-related shift in neural activity may correlate with the shift in navigational strategies reported by many authors. For example, a shift from an allocentric (thought to be hippocampal-dependent) to an egocentric navigational strategy with age has been observed (Rodgers, Sindone, & Moffat, 2012), whereby a flexible use of cues and cognitive mapping in younger adults makes way for a greater route-following and response strategy with an older group (Harris & Wolbers, 2014). However, before testing age difference in navigational strategies or neural responses, we first need to examine if NavWell is sensitive to age-related differences in general. As such, we hypothesized that, in line with previous research, older adults should perform significantly worse during both spatial acquisition and retention. Older adults should take longer to find the invisible target during learning and also search significantly less in the target region upon retest.

Participants Twenty-nine participants were recruited for this experiment. A sample of young adults ($n = 14$), between 19 and 27 years of age (mean age = 22.1 ± 0.6), were recruited from students at Maynooth University and their friends and family. Healthy older adults were likewise recruited from among friends and family and were 55+ years of age (ranging from 55 to 73 years; mean age = 63.8 ± 1.6). All participants tested had normal or corrected-to-normal vision, and both groups were matched for sex. Any participant on medication, had a history of psychiatric, other illness (history of stroke, epilepsy, cardiovascular disease, etc.), history of drug or alcohol abuse was excluded from the study. To ensure that the older adults were cognitively matched with younger adults, we examined both groups on the Montreal

Cognitive Assessment (MoCA). No significant difference was noted between the groups [$t(27) = 1.205$, $p = .239$], with older adults scoring a mean of 27.26 ± 0.8 (out of 30), and the younger group obtaining a mean score of 28.3 ± 0.3. All participants were informed as to the nature of the experiment, provided consent and were told that they were free to withdraw at any stage. The study was approved by Maynooth University ethics committee.

Apparatus The virtual maze for both groups consisted of a medium circular environment (see the details above), again with the platform hidden in the northeast quadrant. Two cues (lights of 50% luminance) were used, which were located in the northeast and northwest positions of the arena (see Fig. 7a). The size of the platform was 15% of the arena area and consisted of a bright green square that became visible when the participant crossed it.
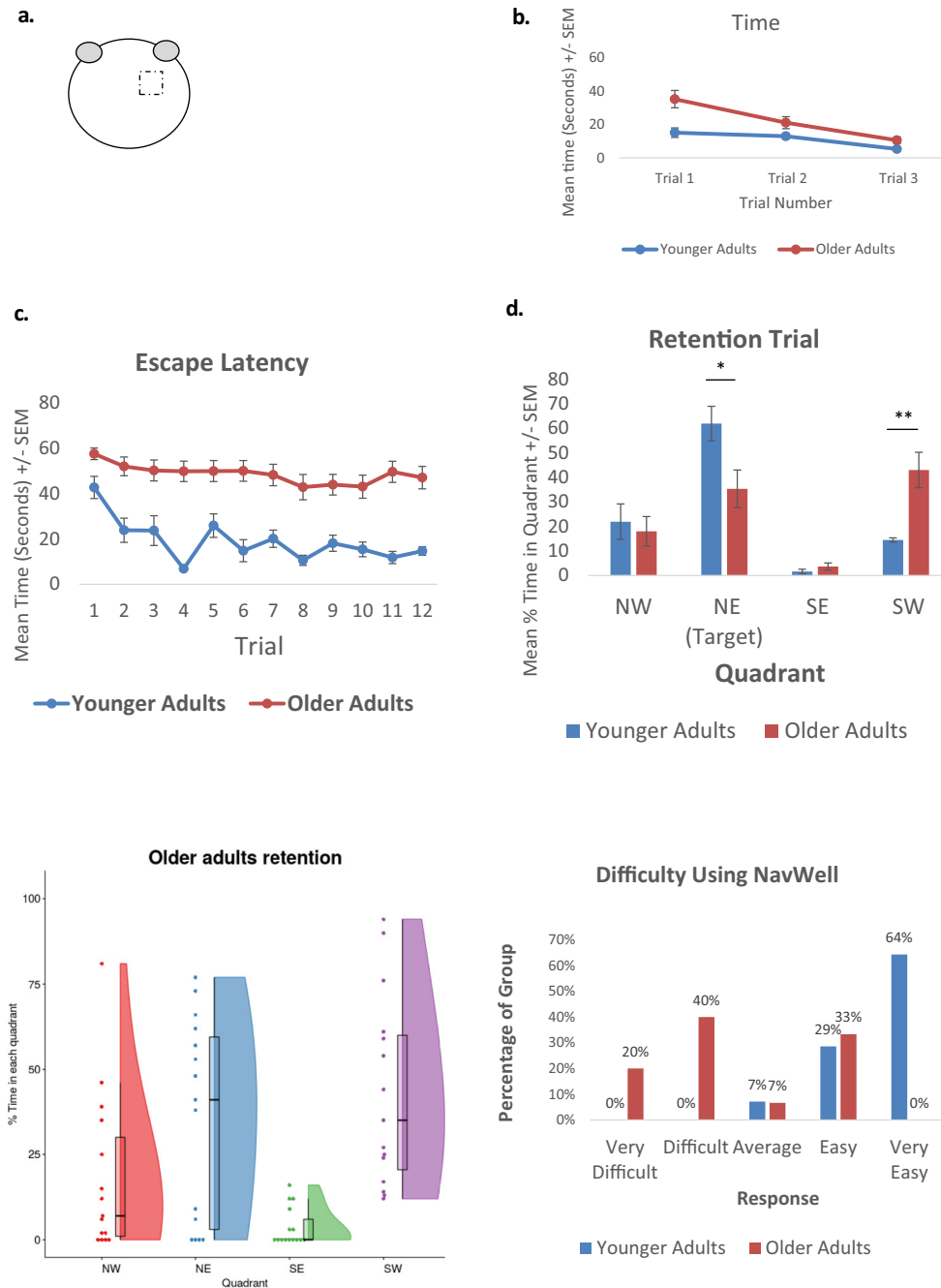

Fig. 7 (a) Illustration of the arena setup for Experiment 4 (young vs. older adults), with the dash–dotted square indicating the location of the platform, and the gray circles representing the location of the cues. (b) Line chart comparing young and older adults on the visible platform task. (c) Line chart comparing spatial learning in young and older adults. (d) Comparison of the time spent searching in various quadrants during the retention trial. (e) Scatterplot showing individual scores for the retention trial (older adults), with boxplots showing the median (horizontal lines) and upper and lower interquartile range. (f) Responses of young and older adults to the ease of use of NavWell

Procedure As in the previous experiments, each participant was given 60 s to find the target in the NavWell task. Once it was found, participants were given 10 s to rest and were instructed to look around the environment and note the location of the cues. Participants were required to complete 12 trials in total. For each trial, the participant started in one of four random locations (north, south, east, or west position). As in the previous experiments, following the acquisition phase, participants were given pen-and-paper control tasks (see above, the NART, TMT, and RAVLT), which took approximately 15–20 min. Following this, the participants were given a single 60-s probe/recall trial in the virtual maze. To ensure that all participants, particularly the older adult group, could use the NavWell controls and had no motor or visual issues, we gave all participants (before training commenced) at least three visible trials. The platform remained lit during these trials; there were no cues in the arena. Participants simply had to make their way to the visible target from random starting positions. Finally, we asked each participant to rate their own spatial abilities and the ease of use of the NavWell tool.

Results

We initially compared the younger and older adults on the various control tasks. No significant difference was found for any measure. The mean number of errors obtained on the NART for the young group was 13.13 ± 2.4, as compared to 16.2 ± 1 for the older group ($p = .269$). The times taken to complete the TMT(B–A) test were 22.2 ± 4.6 s for the young group and 28.5 ± 4.7 s for the older group ($p = .269$). Scores on the RAVLT were also comparable: 51.5 ± 2.4 items recalled for the young group and 44.8 ± 3.8 items for the older group ($p = .138$). We then compared the two groups on how quickly they reached the visible target in NavWell. Both groups had little difficulty navigating toward the target, with a significant effect of trial for both young [$F(2, 26) = 16.65, p = .001$, effect size $= 0.562$] and older [$F(2, 28) = 15.84, p = .001$, effect size $= 0.531$] adults. However, older adults were significantly slower to reach the target [$F(1, 27) = 11.32, p = .0012$, effect size $= 0.295$] than the younger group. Older adults took a mean of 10.5 ± 2 s, as compared to 5.5 ± 0.4 s for the younger group, on the last trial (Fig. 7b).

A 2 (group) × 12 (trial) mixed factorial ANOVA was used to compare how quickly the two groups found the invisible target across the acquisition phase. An overall significant effect for trial [$F(11, 297) = 5.698, p = .001$, effect size $= 0.174$] and a large group effect [$F(1, 27) = 70.06, p = .001$, effect size $= 0.722$] were found. A small significant Trial × Group interaction was noted [$F(11, 297) = 1.854, p = .045$, effect size $= 0.064$]. A follow-up repeated measures ANOVA was conducted for the young group, which showed a large significant effect for trial [$F(11, 143) = 7.298, p = .001$, effect size $= 0.36$]; Bonferroni-corrected t tests showed that young participants were significantly faster at finding the target on Trials 11

(11.7 ± 2.6 s) and 12 (14.64 ± 1.8 s) than on the first trial (42.6 ± 4.8 s). This suggests that younger adults had little difficulty learning the task. However, older adults had greater difficulty and showed no significant effect for trial [$F(11, 154) = 1.100, p = .365$, effect size $= 0.073$; Fig. 7c]. We then compared the distances taken by both groups to reach the target across the acquisition phase. A 2 × 12 mixed factorial ANOVA demonstrated a significant effect for trial [$F(11, 297) = 6.769, p = .001$, effect size $= 0.20$] and a significant group effect [$F(1, 27) = 13.197, p = .001$, effect size $= 0.328$]. In addition, a significant Trial × Group interaction was noted [$F(11, 297) = 3.317, p = .001$, effect size $= 0.109$]. A repeated measures ANOVA was conducted for the young group, and this revealed a significant effect of trial [$F(11, 143) = 7.556, p = .001$, effect size $= 0.368$]. Younger participants took a shorter path on the final trial of learning (64.7 ± 3.4 vm) than on the first trial (180.2 ± 19.8 vm, $p < .001$, Bonferroni-corrected). Although an overall significant effect was shown for the older adults [$F(11, 154) = 2.500, p = .006$, effect size $= 0.152$], Bonferroni-corrected t tests did not show any significant differences. Older adults moved mean distances of 107 ± 17 vm and 157.6 ± 20 vm on Trials 11 and 12, respectively, as compared to 131.4 ± 16 vm on Trial 1 (see Fig. 7c).

Retention was assessed by allowing both groups 60 s to search the arena for the target (without it being present). A 2 (group) × 4 (quadrant) mixed factorial ANOVA was conducted. We found a significant effect for quadrant [$F(3, 81) = 17.11, p = .001$, effect size $= 0.388$] and a Group × Quadrant interaction [$F(3, 81) = 6.001, p = .001$, effect size $= 0.182$]. Follow-up analyses revealed that the young group searched significantly more in the target area [NE quadrant, $t(27) = 2.548, p = .017$] than did the older group. In contrast, the older group spent significantly more time in the starting quadrant [SW, $t(27) = -3.781, p = .001$] than the younger group (see Fig. 7d). Interestingly, when we plotted the individual scores for just the older adult group (Fig. 7e), we saw that many older adults recalled the target very well (searching in the correct target more than 50% of the time), whereas others performed poorly.

From the questionnaire, 40% of older adults found the task average or easy (Fig. 7f), while 40% found it difficult or very difficult (20%). This is in contrast to the younger adults, where 93% found the task easy or very easy. In response to the question of how participants rated their own spatial ability, older adults showed a spread of responses: very poor (7%), poor (20%), average (20%), good (20%), or very good (33%). In contrast, 7% of younger adults rated their own abilities as very good, 57% as good, 21% as average, and 14% as poor (these data are not shown).

Discussion

Older adults, as in previous research (Head & Isom, 2010; Moffat & Resnick, 2002), were generally slower to find the

hidden target and covered a larger distance compared to younger adults. While the distance travelled decreased across trials for older participants, suggesting that they learned the task to a degree, this decrease was inconsistent. Furthermore, examination of the retention trial show that older adults searched both in the target NE quadrant and in the SW quadrant, the starting point of the trial. This is in contrast to the younger adults, who searched primarily in the target quadrant. These results perhaps suggest that some older adults needed more time at the start before deciding on where to search. The need for more time to search was also reflected in the visible platform task. Although older adults seemed to be able to manage the tool relatively well and navigated toward the visible target, they were significantly slower than the younger participants. It took the older group an average of 10 s to reach the target compared to 5 s for the younger participants by Trial 3. Slower reaction times are not just confined to spatial navigation but are observed across a range of cognitive domains (see Walshe, Patterson, Commins, & Roche, 2015).

Alternatively, when looking at individual performances of older adults, particularly during the retention trial, we observe a large variability in response. For example, although the majority of older adults (9/15) searched correctly for the target in the NE quadrant (spending more than 50% of the time searching here), a number of older adults (6/15) showed very poor recall. This variability in performance has also been observed across a number of cognitive domains, variability increases with age (Bielak, Cherbuin, Bunce, & Anstey, 2014) and is often associated with maladaptive traits including poorer cognition and everyday functioning (Burton, Strauss, Hultsch, & Hunter, 2009). Indeed, such variability may help predict age-related diseases (e.g., Holtzer, Jacobs, & Demetriou, 2019). Another issue that arose from the current study is that many older adults did not rate their own spatial abilities highly and importantly found the task quite difficult. This is important, and in future iterations of the task we will endeavor to make NavWell easier for all future participants, particularly those who are older or have age-related disorders.

In sum, this pilot study suggests that the poorer performance observed in older adults may be due to a number of factors including a poorer spatial ability, a greater performance variability, a slower reaction time and a greater difficulty in using the tool. Each of these factors will need to be examined further and separated in order to understand the full complexity of age-related navigation.

General discussion

We have developed a free, easy-to-use navigation tool for use primarily by researchers but possibly by clinicians and educators. The task requires participants to navigate an arena to find a hidden target and recall this location on subsequent trials using various cues in the environment. For researchers the task offers a simpler alternative than currently available software, allowing for the design of experiments in a simple fashion. Researchers can select the shape and size of the environment, select various cues and set up various experimental protocols. For educators or clinicians, the task has a preprepared environmental setup and experimental protocol and is ready for immediate use. This makes it easy to apply in a class-room setting or possibly in the clinic using procedures that have been validated and tested.

Virtual reality tools are extremely useful for research, especially those that allow participants to be fully immersive. However, researchers should remain cautious and be aware of their limitations. Such limitations may be equally applied to NavWell, including possible motion sickness, expensive equipment and the practical use of HMDs with children, older adults and other vulnerable populations (see the discussion above). Furthermore, the question of how translatable the VR environment is to a real world situation should be considered. Indeed, many authors have argued that navigation in small, virtual environments may not be comparable to natural large scale navigation (Learmonth, Newcombe, Sheridan, & Jones, 2008, but see Claessen, Visser-Meily, de Rooij, Postma, & van der Ham, 2016, for counterargument). As gaming becomes increasingly popular, participants should also be assessed in terms of their familiarity with computers, gaming and VR technology. Familiarity with technology and frequency of use should all be considered before commencing a VR experiment, possibly using questionnaires.

The experiments presented here have tested reference memory, as this has been the primary focus of many behavioral (Padilla, Creem-Regehr, Stefanucci, & Cashdan, 2017), clinical (Schneider et al., 2017), pharmacological (Farina & Commins, 2016), and ageing (Zhong et al., 2017) studies, in both humans and animals. The virtual water maze task and NavWell in particular, however, can be used to test a range of other cognitive functions. For example, the observed shift in older adults toward a more egocentric strategy and away from an allocentric navigational strategy (Bohbot et al., 2012) can easily be accommodated by NavWell by simply comparing the ability of older adults to perform a task in which the platform is visible (egocentric/beacon) or invisible (allocentric/ multiple distal cues). Spatial working memory, a key issue with frontal lobe patients (Tsuchida & Fellows, 2013) and people with schizophrenia (Spieker, Astur, West, Griego, & Rowland, 2012), can also be examined with NavWell by moving the invisible platform's position between trial blocks and by keeping its position constant within a trial block (see Nakazawa, McHugh, Wilson, & Tonegawa, 2004). Likewise, procedural memory can be tested by adopting the methods used by Kealy et al. (2008) or Shires and Aggleton (2008) in the animal literature, whereby both the starting position and the goal always remains fixed. Although briefly

discussed with Experiment 3, the shape of environment may play an important role in spatial reorientation and this could be explored in greater depth using NavWell. Nonmnemonic issues such as motor dexterity, motivation and visual functioning can also be examined with NavWell, particularly using the visible platform option. As NavWell develops, other features including the expansion of the current cues, addition of intramaze cues, replaying of movements, deeper analyses of these movements will be added, which will allow researchers to explore spatial behavior in greater depth. Ease of use will be an important consideration for use with older and more vulnerable populations.

Despite the versatility of the virtual water maze task, one of the major difficulties with conducting research is that many virtual environments are not standardized. Furthermore, there is a lack of standardization of experimental protocols across research groups. This has implications for the replication of results and perhaps explains why the task is not more readily available as a testing tool in the clinic. For example, many groups create their own virtual environments that range from using a circular pool within a larger room (of various sizes and shapes, see, e.g., Skelton et al., 2006) to virtual islands with buried treasure (Schoenfeld, Foreman, & Leplow, 2014). The surrounding landmarks used to test reference memory also vary in size, type, and number. Some groups use two cues on the circumference of the pool, others use four cues located on the walls of the room and others have windows on the walls that open out to different vistas (Herting & Nagel, 2012; Livingstone & Skelton, 2007; Nedelska et al., 2012). Likewise, the relative size of the invisible target to the arena, if reported, varies considerably from 2% to 17% (Goodrich-Hunsaker, Livingstone, Skelton, & Hopkins, 2010; Schautzer, Hamilton, Kalla, Strupp, & Brandt, 2003). Even the number of trials, number of starting positions and the time allowed to find the platform differ across researchers; for example, protocols for the invisible platform condition include 10 × 50 s (Schautzer et al., 2003), 16 × 74 s (Goodrich-Hunsaker et al., 2010), and 7 × 3 min (Kallai, Makany, Karadi, & Jacobs, 2005) trials. Such variation may allow some participants to find the platform much easier compared to others. Because latencies and dwell times in various regions can not be directly compared across research groups, this makes the reproduction and replication of results very difficult to achieve. To overcome these issues when working with NavWell, we have included (in the supplementary information) a recommended method and set of procedures to be used when testing spatial navigation. These recommendations are based on our own experiments and on reference to the literature. For researchers who wish to use NavWell for testing other mnemonic and nonmnemonic functions (aside from reference memory), we have included additional different protocols for testing general perceptual and motor functioning (visible/cued platform condition) and spatial working memory.

Compliance with ethical standards

References

Antonova, E., Parslow, D., Brammer, M., Dawson, G. R., Jackson, S. H., & Morris, R. G. (2009). Age-related neural activity during allocentric spatial memory. *Memory, 17*, 125–143.

Artigas, A. A., Aznar-Casanova, J. A., & Chamizo, V. D. (2005). Effects of absolute proximity between landmark and platform in a virtual Morris pool task with humans. *International Journal of Computational Psychology, 18*, 224–238.

Astur, R. S., Taylor, L. B., Mamelak, A. N., Philpott, L., & Sutherland, R. J. (2002). Humans with hippocampus damage display severe spatial memory impairments in a virtual Morris water task. *Behavioural Brain Research, 132*, 77–84.

Astur, R. S., Tropp, J., Sava, S., Constable, R. T., & Markus, E. J. (2004). Sex differences and correlations in a virtual Morris water task, a virtual radial arm maze, and mental rotation. *Behavioural Brain Research, 151*, 103–115.

Baeta-Corral, R., & Giménez-Llort, L. (2015). Persistent hyperactivity and distinctive strategy features in the Morris water maze in 3xTg-AD mice at advanced stages of disease. *Behavioral Neuroscience, 129*, 129–137.

Barry, D. N., Coogan, A. N., & Commins, S. (2016). The time course of systems consolidation of spatial memory from recent to remote retention: A comparison of the immediate early genes *Zif268, c-Fos* and *Arc. Neurobiology of Learning and Memory, 128*, 46–55. https://doi.org/10.1016/j.nlm.2015.12.010

Bielak, A. A., Cherbuin, N., Bunce, D., & Anstey, K. J. (2014). Intraindividual variability is a fundamental phenomenon of aging: Evidence from an 8-year longitudinal study across young, middle, and older adulthood. *Developmental Psychology, 50*, 143–151.

Bliss, T. V., & Lomo, T. (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *Journal of Physiology, 232*, 331–356.

Bohbot, V. D., McKenzie, S., Konishi, K., Fouquet, C., Kurdi, V., Schachar, R., . . . Robaey, P. (2012). Virtual navigation strategies from childhood to senescence: Evidence for changes across the life span. *Frontiers in Aging Neuroscience, 4*, 28. https://doi.org/10.3389/fnagi.2012.00028

Burton, C. L., Strauss, E., Hultsch, D. F., & Hunter, M. A. (2009). The relationship between everyday problem solving and inconsistency in reaction time in older adults. *Aging, Neuropsychology and Cognition, 16*, 607–632.

Chamizo, V. D., & Rodrigo, T. (2004). Effect of absolute spatial proximity between a landmark and a goal. *Learning and Motivation, 35*, 102–114.

Cheng, K. (1986). A purely geometric module in the rat's spatial representation. *Cognition, 23*, 149–178.

Claessen, M. H., Visser-Meily, J. M., de Rooij, N. K., Postma, A., & van der Ham, I. J. (2016). A direct comparison of real-world and virtual navigation performance in chronic stroke patients. *Journal of the International Neuropsychological Society, 22*, 467–477.

Commins, S., & Fey, D. (2019). Understanding the role of distance, direction and cue salience in an associative model of landmark learning. *Scientific Reports, 9*, 2026.

Coughlan, G., Laczó, J., Hort, J., & Hornberger, M. (2018). Spatial navigation deficits: Overlooked cognitive marker for preclinical Alzheimer disease? *Nature Reviews Neurology, 14*, 496–506.

Craig, S., & Commins, S. (2007). Plastic and metaplastic changes in the CA1 and subicular projections to the entorhinal cortex. *Brain Research*, *1147*, 124–139.

Diviney, M., Fey, D., & Commins, S. (2013). Hippocampal contribution to vector model hypothesis during cue-dependent navigation. *Learning and Memory*, *20*, 367–378.

Farina, F. R., Burke, T., Coyle, D., Jeter, K., McGee, M., O'Connell, J., . . . Commins S. (2015). Learning efficiency: The influence of cue salience during spatial navigation. *Behavioral Processes*, *116*, 17–27.

Farina, F. R., & Commins, S. (2016). Differential expression of immediate early genes *Zif268* and *c-Fos* in the hippocampus and prefrontal cortex following spatial learning and glutamate receptor antagonism. *Behavioural Brain Research*, *307*, 194–198.

Folley, B. S., Astur, R., Jagannathan, K., Calhoun, V. D., & Pearlson, G. D. (2010). Anomalous neural circuit function in schizophrenia during a virtual Morris water task. *NeuroImage*, *49*, 3373–3384.

Fu, E., Bravo, M., & Roskos, B. (2015). Single-destination navigation in a multiple-destination environment: A new "later-destination attractor" bias in route choice. *Memory & Cognition*, *43*, 1043–1055. https://doi.org/10.3758/s13421-015-0521-7

Goodrich-Hunsaker, N. J., Livingstone, S. A., Skelton, R. W., & Hopkins, R. O. (2010). Spatial deficits in a virtual water maze in amnesic participants with hippocampal damage. *Hippocampus*, *20*, 481–491.

Grieves, R. M., & Jeffery, K. J. (2017). The representation of space in the brain. *Behavioral Processes*, *135*, 113–131.

Hamilton, D. A., Driscoll, I., & Sutherland, R. J. (2002). Human place learning in a virtual Morris water task: Some important constraints on the flexibility of place navigation. *Behavioural Brain Research*, *129*, 159–170.

Harris, M. A., & Wolbers, T. (2014). How age-related strategy switching deficits affect wayfinding in complex environments. *Neurobiology of Aging*, *35*, 1095–1102.

Hartley, T., Maguire, E. A., Spiers, H. J., & Burgess, N. (2003). The well-worn route and the path less traveled: Distinct neural bases of route following and wayfinding in humans. *Neuron*, *37*, 877–888.

Harvey, D. R., Brant, L., & Commins S. (2009). Differences in cue-dependent spatial navigation may be revealed by in-depth swimming analysis. *Behavioral Processes*, *82*, 190–197.

Hayward, A., Good, M. A., & Pearce, J. M. (2004). Failure of a landmark to restrict spatial learning based on the shape of the environment. *Quarterly Journal of Experimental Psychology*, *57B*, 289–314.

Head, D., & Isom, M. (2010). Age effects on wayfinding and route learning skills. *Behavioural Brain Research*, *209*, 49–58.

Hermer, L., & Spelke, E. (1996). Modularity and development: The case of spatial reorientation. *Cognition*, *61*, 195–232.

Hermer, L., & Spelke, E. S. (1994). A geometric process for spatial reorientation in young children. *Nature*, *370*, 57–59.

Herting, M. M., & Nagel, B. J. (2012). Aerobic fitness relates to learning on a virtual Morris water task and hippocampal volume in adolescents. *Behavioural Brain Research*, *233*, 517–525.

Holtzer, R., Jacobs, S., & Demetriou, E. (2019). Intraindividual variability in verbal fluency performance is moderated by and predictive of mild cognitive impairments. *Neuropsychology*. Advance online publication. https://doi.org/10.1037/neu0000576

Hupont, I., Gracia, J., Sanagustín, L., & Gracia, M. A. (2015). How do new visual immersive systems influence gaming QoE? A use case of serious gaming with Oculus Rift. In *Seventh International Workshop on Quality of Multimedia Experience (QoMEX)* (pp. 1–6). Piscataway, NJ: IEEE Press. https://doi.org/10.1109/QoMEX.2015.7148110

Jansen-Osmann, P., & Fuchs, P. (2006). Wayfinding behavior and spatial knowledge of adults and children in a virtual environment: The role of landmarks. *Experimental Psychology*, *53*, 171–181.

Kallai, J., Makany, T., Karadi, K., & Jacobs, W. J. (2005). Spatial orientation strategies in Morris-type virtual water task for humans. *Behavioural Brain Research*, *159*, 187–196.

Kargar, M., Askari, S., Khoshaman, A., & Mohammadi, A. (2019). Differential diagnosis of schizophrenia and schizoaffective disorder from normal subjects using virtual reality. *Psychiatry Research*, *273*, 378–386.

Kealy, J., Diviney, M., Kehoe, E., McGonagle, V., O'Shea, A., Harvey, D., & Commins, S. (2008). The effects of overtraining in the Morris water maze on allocentric and egocentric learning strategies in rats. *Behavioural Brain Research*, *192*, 259–263. https://doi.org/10.1016/j.bbr.2008.04.009

Keith, J. R., Priester, C., Ferguson, M., Salling, M., & Hancock, A. (2008). Persistent increases in the pool of doublecortin-expressing neurons in the hippocampus following spatial navigation training. *Behavioural Brain Research*, *188*, 391–397.

Kennedy, R. S., Lane, N. E., Berbaum, K. S., & Lilienthal, M. G. (1993). Simulator sickness questionnaire: An enhanced method for quantifying simulator sickness. *International Journal of Aviation Psychology*, *3*, 203–220.

Kolarik, B. S., Shahlaie, K., Hassan, A., Borders, A. A., Kaufman, K. C., Gurkoff, G., … Ekstrom A. D. (2016). Impairments in precision, rather than spatial strategy, characterize performance on the virtual Morris Water Maze: A case study. *Neuropsychologia*, *80*, 90–101.

Learmonth, A. E., Newcombe, N. S., Sheridan, N., & Jones, M. (2008). Why size counts: Children's spatial reorientation in large and small enclosures. *Developmental Science*, *11*, 414–426.

Lester, A. W., Moffat, S. D., Wiener, J. M., Barnes, C. A., & Wolbers, T. (2017). The aging navigational system. *Neuron*, *95*, 1019–1035.

Lew, A. R., Usherwood, B., Fragkioudaki, F., Koukoumi, V., Smith, S. P., Austen, J. M., & McGregor, A. (2014). Transfer of spatial search between environments in human adults and young children (*Homo sapiens*): Implications for representation of local geometry by spatial systems. *Developmental Psychobiology*, *56*, 421–434.

Livingstone, S. A., & Skelton, R. W. (2007). Virtual environment navigation tasks and the assessment of cognitive deficits in individuals with brain injury. *Behavioural Brain Research*, *185*, 21–31.

Lubyk, D. M., Dupuis, B., Gutiérrez, L., & Spetch, M. L. (2012). Geometric orientation by humans: Angles weigh in. *Psychonomic Bulletin & Review*, *19*, 436–442. https://doi.org/10.3758/s13423-012-0232-z

Moffat, S. D., Elkins, W., & Resnick, S. M. (2006). Age differences in the neural systems supporting human allocentric spatial navigation. *Neurobiology of Aging*, *27*, 965–972.

Moffat, S. D., & Resnick, S. M. (2002). Effects of age on virtual environment place navigation and allocentric cognitive mapping. *Behavioral Neuroscience*, *116*, 851–859. https://doi.org/10.1037//0735-7044.116.5.851

Morris, R. G. (1981). Spatial localization does not require the presence of local cues. *Learning and Motivation*, *12*, 239–260.

Morris, R. G., Garrud, P., Rawlins, J. N., & O'Keefe, J. (1982). Place navigation impaired in rats with hippocampal lesions. *Nature*, *297*, 681–683.

Mosadeghi, S., Reid, M. W., Martinez, B., Rosen, B. T., & Spiegel, B. M. (2016). Feasibility of an immersive virtual reality intervention for hospitalized patients: An observational cohort study. *JMIR Mental Health*, *3*, e28.

Munafo, J., Diedrick, M., & Stoffregen, T. A. (2017). The virtual reality head-mounted display Oculus Rift induces motion sickness and is sexist in its effects. *Experimental Brain Research*, *235*, 889–901. https://doi.org/10.1007/s00221-016-4846-7

Nakazawa, K., McHugh, T. J., Wilson, M. A., & Tonegawa, S. (2004). NMDA receptors, place cells and hippocampal spatial memory. *Nature Reviews Neuroscience*, *5*, 361–372.

Nedelska, Z., Andel, R., Laczó, J., Vlcek, K., Horinek, D., Lisy, J., … Hort, J. (2012). Spatial navigation impairment is proportional to

right hippocampal volume. *Proceedings of the National Academy of Sciences*, *109*, 2590–2594.

Nelson, H. E. (1982). National Adult Reading Test (NART): Test manual. Windsor: NFER.

Newhouse, P., Newhouse, C., & Astur, R. S. (2007). Sex differences in visual-spatial learning using a virtual water maze in pre-pubertal children. *Behavioural Brain Research*, *183*, 1–7.

O'Keefe, J., & Nadel, L. (1979). Précis of O'Keefe and Nadel's *The hippocampus as a cognitive map. Behavioral and Brain Sciences*, *2*, 487–533. https://doi.org/10.1017/S0140525X00063949

Padilla, L. M., Creem-Regehr, S. H., Stefanucci, J. K., & Cashdan, E. A. (2017). Sex differences in virtual navigation influenced by scale and navigation experience. *Psychonomic Bulletin & Review*, *24*, 582–590.

Pearce, J. M. (2009). The 36th Sir Frederick Bartlett lecture: An associative analysis of spatial learning. *Quarterly Journal of Experimental Psychology*, *62*, 1665–1684. https://doi.org/10.1080/17470210902805589

Pearce, J. M., Good, M. A., Jones, P. M., & McGregor, A. (2004). Transfer of spatial behavior between different environments: Implications for theories of spatial learning and for the role of the hippocampus in spatial learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *30*, 135–147.

Ponto, K., Chen, K., Tredinnick, R., & Radwin, R. G. (2014). Assessing exertions: How an increased level of immersion unwittingly leads to more natural behavior. In *2014 IEEE Annual International Symposium on Virtual Reality* (pp. 107–108). Piscataway, NJ: IEEE Press. https://doi.org/10.1109/vr.2014.6802074

Possin, K. L., Sanchez, P. E., Anderson-Bergman, C., Fernandez, R., Kerchner, G. A., Johnson, E. T., … Finkbeiner, S. (2016). Cross-species translation of the Morris maze for Alzheimer's disease. *Journal of Clinical Investigation*, *126*, 779–783.

Redhead, E. S., & Hamilton, D. A. (2009). Evidence of blocking with geometric cues in a virtual watermaze. *Learning and Motivation*, *40*, 15–34.

Reiten, R. M. (1958). Validity of the Trail Making Test as an indicator of organic brain damage. *Perceptual & Motor Skills*, *8*, 271–276.

Rey, A. (1941). L'examen psychologique dans les cas d'encéphalopathie traumatique (Les problems). *Archives de Psychologie*, *28*, 215–285.

Rodgers, M. K., Sindone, J. A., 3rd, & Moffat, S. D. (2012). Effects of age on navigation strategy. *Neurobiology of Aging*, *33*, 202.e15–22.

Rodrigo, T. (2002). Navigational strategies and models. *Psicologica*, *23*, 3–32.

Santos, B. S., Dias, P., Pimentel, A., Baggerman, J. W., Ferreira, C., Silva, S., & Madeira, J. (2009). Head-mounted display versus desktop for 3D navigation in virtual reality: A user study. *Multimedia Tools and Applications*, *41*, 161–181. https://doi.org/10.1007/s11042-008-0223-2

Schautzer, F., Hamilton, D., Kalla, R., Strupp, M., & Brandt, T. (2003). Spatial memory deficits in patients with chronic bilateral vestibular failure. *Annals of the New York Academy of Sciences*, *1004*, 316–324.

Schneider, C. B., Linse, K., Schönfeld, R., Brown, S., Koch, R., Reichmann, H., … Storch, A. (2017). Spatial learning deficits in Parkinson's disease with and without mild cognitive impairment. *Parkinsonism and Related Disorders*, *36*, 83–88.

Schoenfeld, R., Foreman, N., & Leplow, B. (2014). Ageing and spatial reversal learning in humans: Findings from a virtual water maze. *Behavioural Brain Research*, *270*, 47–55.

Shires, K. L., & Aggleton, J. P (2008). Mapping immediate-early gene activity in the rat after place learning in a water-maze: The importance of matched control conditions. *European Journal of Neuroscience*, *28*, 982–996.

Skelton, R. W., Ross, S. P., Nerad, L., & Livingstone, S. A. (2006). Human spatial navigation deficits after traumatic brain injury shown in the arena maze, a virtual Morris water maze. *Brain Injury*, *20*, 189–203.

Spieker, E. A., Astur, R. S., West, J. T., Griego, J. A., & Rowland, L. M. (2012). Spatial memory deficits in a virtual reality eight-arm radial maze in schizophrenia. *Schizophrenia Research*, *135*, 84–89.

Stankiewicz, B. J., Legge, G. E., Mansfield, J. S., & Schlicht, E. J. (2006). Lost in virtual space: Studies in human and ideal spatial navigation. *Journal of Experimental Psychology: Human Perception and Performance*, *32*, 688–704. https://doi.org/10.1037/0096-1523.32.3.688

Takahashi, N., Kawamura, M., Shiota, J., Kasahata, N., & Hirayama, K. (1997). Pure topographic disorientation due to right retrosplenial lesion. *Neurology*, *49*, 464–469.

Tsuchida, A., & Fellows, L. K. (2013). Are core component processes of executive function dissociable within the frontal lobes? Evidence from humans with focal prefrontal damage. *Cortex*, *49*, 1790–1800.

Walshe, E. A., Patterson, M. R., Commins, S., & Roche, R. A. (2015). Dual-task and electrophysiological markers of executive cognitive processing in older adult gait and fall-risk. *Frontiers in Human Neuroscience*, *9*, 200. https://doi.org/10.3389/fnhum.2015.00200

Witmer, B. G., & Singer, M. J. (1994). Measuring immersion in virtual environments (ARI Technical Report 1014). Alexandria: U.S. Army Research Institute for the Behavioral and Social Sciences.

Zanco, M., Plácido, J., Marinho, V., Ferreira, J. V., de Oliveira, F., Monteiro-Junior, R., … Deslandes, A. (2018). Spatial navigation in the elderly with Alzheimer's disease: A cross-sectional study. *Journal of Alzheimer's Disease*, *66*, 1683–1694.

Zhong, J. Y., Magnusson, K. R., Swarts, M. E., Clendinen, C. A., Reynolds, N. C., & Moffat, S. D. (2017). The application of a rodent-based Morris water maze (MWM) protocol to an investigation of age-related differences in human spatial learning. *Behavioural Neuroscience*, *131*, 470–482.