Explaining World-Wide Variation in Navigation Ability from Millions of People: Citizen Science Project Sea Hero Quest

Spiers, H. J.[1] , Coutrot, A.[2,3] , Hornberger, M.[4]

1. Institute of Behavioural Neuroscience, Department of Experimental Psychology, Division of Psychology and Language Sciences, University College London, London, UK.
2. Laboratoire des Sciences du Numérique de Nantes, CNRS, Nantes, France.
3. Laboratoire d'Informatique en Image et Systèmes d'information, CNRS, Lyon, France.
4. Norwich Medical School, University of East Anglia, Norwich, UK.

Author email addresses: h.spiers@ucl.ac.uk, antoine.coutrot@ls2n.fr, M.Hornberger@uea.ac.uk

Key words: Dementia, Wayfinding, Culture, Spatial Cognition, Gender, Aging, Games

Abstract

Navigation ability varies widely across humans. Prior studies have reported that being younger and male has an advantage for navigation ability. However, these studies have generally involved small numbers of participants from a handful of western countries. Here we review findings from our project Sea Hero Quest which used a video game for mobile and tablet devices to test 3.9 million people on their navigation ability, sampling across every nation state and from 18-99 years of age. Results revealed that the task has good ecological validity and, across all countries sufficiently sampled (N=63), age is linked to a near linear decline in navigation ability from the early 20s. All countries showed a male advantage, but this varied considerably and could be partly predicted by gender inequality. We found that those reported growing up in a city were on average worse at navigating than those who grew up outside cities, and that navigation performance helped identify those at greater genetic-risk of Alzheimer's disease. We discuss the advantages and challenges of using a mobile app to study cognition and the future avenues for understanding individual differences in navigation ability arising from this research.

Introduction

Some humans are able to rapidly learn the layout of a new place and others can become lost easily (Wolbers & Hegarty, 2010; Weisberg & Newcombe 2016; Ekstrom et al., 2018; Burles & Iaria, 2020). Understanding this variation in behaviour is important, not only for providing insights into cognition and brain function, but for the capacity to detect and monitor the progression of Alzheimer’s disease (AD), where deficits in spatial memory appear early in the disease (Coughlan et al., 2018a). AD affects over 50 million people world-wide and its prevalence is increasing in countries with an aging population: this number will more than triple by 2050 (Alzheimer’s Disease International, 2018). Thus having the capacity to detect and monitor the disease is important, particularly if there is the future possibility of treatment requiring early interventions. Moreover, given the challenges of creating a test that is widely available and not culturally affected, such as those that use verbal material, this format is particularly attractive.

While developing a test of spatial navigation has been of interest, there have been surprisingly few advances. Spatial navigation is not an ability that can easily be tested in standard testing rooms used for clinical assessment as these cannot be controlled or manipulated sufficiently for standardised assessment. Thus, virtual reality (VR) is a core tool that has been turned to solve this problem (Coughlan et al., 2018a). Given that AD prevalence increases with age, the use of VR raises challenges because the older aged population tends to have less familiarity with VR and may struggle with operating controls. This has been a significant challenge in the past with desk-top VR controlled by a joystick or mouse. The recent advent of widespread touch screen technology on tablet and mobile devices has changed that. Such devices provide a much more intuitive appeal to older populations. An additional challenge faced when developing a test of spatial navigation is collecting a sufficient sample large enough to account for the wide variation in human ability. Testing thousands of people on a VR task is time consuming and costly.

Here we review recent research by our team that surmounted the challenge of mass testing by developing a flat screen VR navigation task in a video game app - Sea Hero Quest - on tablet and mobile devices that collected data in every country world-wide, assessing the performance of 3.9 million people (Coutrot et al., 2018; Coutrot et al. 2020). We discuss how this was achieved, what the results have revealed about human spatial navigation and also future directions.

Development of a world-wide virtual assessment of navigation ability: Citizen science on a global scale

The project arose from the intention of Deutsche Telekom (T-mobile) to engage people with a campaign using mobile phones to achieve a positive change in the world. The idea proposed by one of us (MH) of creating a navigation task in an engaging video game that could provide the benchmark assessment of navigation ability was chosen as the project they would support, with HS brought in as co-lead for the scientific development. This led to the engagement of creative agency Saachi and Saachi and the video games company Glitchers Ltd, who created the navigation test embedded in the video game Sea Hero Quest (see Morgan, 2016). Alongside the commercial side of the project, careful guidance and grant support was given throughout the full length of the project by the charity

Alzheimer's Research UK and led one of us (AC) joining the team to analyse the data. Two key design issues needed addressing: 1. What tests would work in a successful video game? 2. How can data be collected remotely from participants playing the video game on their phone?

The idea of testing navigation by steering a boat through oceans and marshes in search of sea creatures came from the game developers the Glitchers Ltd: Max Scott-Slade, Hugo Scott-Slade and Matt Hyde. We found this approach beneficial for two reasons: 1) for many learning to steer a boat in the real world is difficult and this makes the challenge of learning the game controls appear more consistent with the real world, 2) with water it is clear where you can go and where you cannot go. Through discussions with the Glitchers and collaborators, three spatial tasks were selected for ease in a touch screen video game (Fig. 1a). While a logical choice would have been to test memory using a variant of the famous Morris Water Maze, in which the participant must locate a hidden location somewhere in a large circular pool or arena (Morris et al., 1982; Moffat et al., 2001), we opted not to use this because it requires regular circling behaviour in VR. This could potentially cause nausea, and the repetitive experience would be less engaging for participants. Instead we opted to examine wayfinding, path integration and spatial working memory (Fig. 1).

Wayfinding refers to the demand to travel through an environment to a remembered or indicated goal location (Wiener et al., 2009). Wayfinding was chosen as it forms a common part of everyday navigation experience. Remembering the way to the shops, travelling to a gallery to meet friends and navigating an unknown city in search of restaurant are all examples of wayfinding (Ekstrom et al., 2018). Wayfinding is also a task in which numerous studies have examined gender differences and the neural correlates (Ekstrom et al., 2018). To keep the task short for maximum engagement with a large sample, we tested wayfinding by providing participants with a map that displayed check-points to reach in a numbered order which had to be memorised (Fig 1b). The participant had unlimited time to view the map before they began navigation to the checkpoints. Participants were rewarded three stars if they reached the checkpoints quickly, two if slower, and if they took too long, an arrow indicating the direction to the checkpoints appeared and they received one star. The 45 check-point levels created varied in a range of parameters: number of checkpoints, complexity of the environment, the size of the environment, the presence of local and global landmarks, fog, surprise forced detours and, in the later levels, waves. These were included to allow analysis of the impact of environmental features on navigation, but also to vary the game experience.

Path integration is the ability to use self-motion information to determine the distance and direction travelled from a point of origin, and calculate a homing vector to the origin (Wiener et al., 2009). A path integration task was chosen due to prior work suggesting it may help discriminate Alzheimer's disease from other dementias (Tu et al., 2017). Our path integration task (flare levels) required following a path to collect a flare gun and shooting it back to the start of the level, choosing the correct direction (Fig 1d). There were 15 of these levels included in the game. Spatial working memory in our case was memory for a set of locations recently visited in a virtual radial maze (Bohbot et al., 2007). The virtual radial maze helped provide a separate measure of spatial memory but also a measure of strategy use. Five radial mazes were included in the game as bonus levels. In these, participants had to collect hidden stars from the end of 3 of the 6 arms in the radial maze, with 3 arms blocked. After collecting all three stars, all arms were unblocked and the participants had to collect the stars from the arms that had previously been blocked. After solving the task participants were also asked whether they navigated by counting the arms, or used a landmark or a combination.

In addition to the three spatial tasks, every 5th level of the 75 levels in the game included a task in which participants were required to chase and photograph a sea monster. Each of these 15 levels required tapping left or right to move the boat only left or right to avoid oncoming rocks and waves. If successful after a short period the view switched to a view through a camera and the participant was required to capture a photograph for their log in the game menu. This was included to vary the game play and create photographs that could allow users to share progress, helping to disseminate the game via social media. The data from these levels was not recorded to minimize the data being transferred and stored with the app. Because a measure of visuomotor video games skill was measured in the time taken to complete tutorial levels levels 1 and 2, it was decided the extra data from these monster chase levels was not essential.

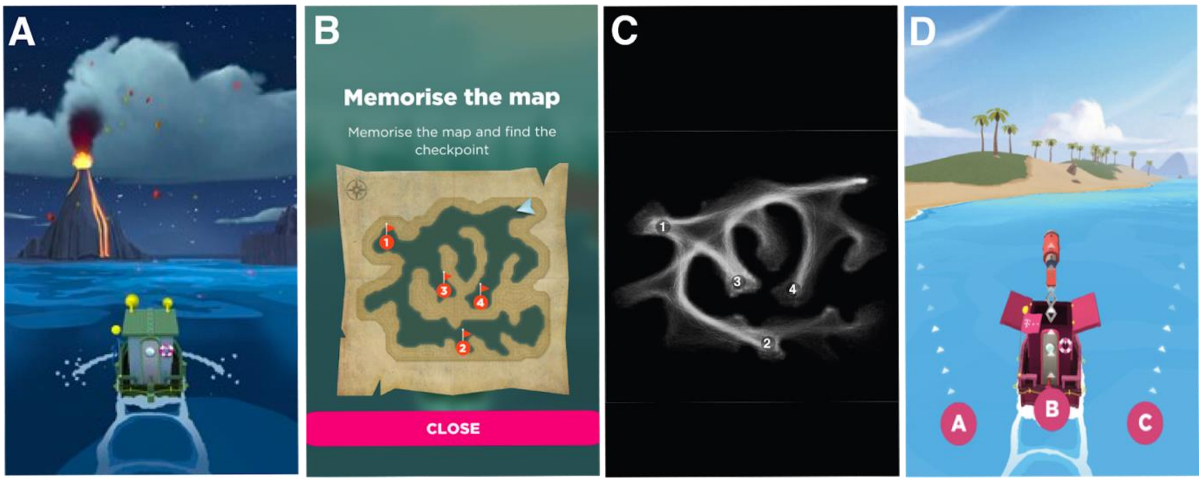

Figure 1 | Task Design - **a-b** Wayfinding task: a map of the level featuring the ordered set of checkpoints to reach is presented and disappears when the game starts. **c -** Superposition of 1,000 individual trajectories randomly sampled from level 32. **d -** Path-integration task: after navigating the level, participants must shoot a flare back to the starting point. Reproduced from Coutrot et al., 2018.

To collect data from millions of participants world-wide it was necessary to optimally collect the data and advertise the game. Data was recorded locally to the participant's device and sent encrypted to a secure server, in this case run by T-mobile. Consent for this was provided by the UCL ethics board and informed consent provided with the game. Several strategies were taken to advertise for participants. Saatchi and Saatchi Ltd played a key role in development of the project in this direction, creating several film and animation adverts about the study with high end production values and a developing emotive story linked to dementia to motivate players to join. The game play was linked to helping a person collect the pages of his father's travel journal that had been lost and needed to be found. The game itself allowed players to share progress and the upgrading of their boat via Facebook and Twitter. T-mobile specifically advertised the game to its millions of customers. Alzheimer's research UK promoted it to its supporters. Finally, the world's most watched youtuber, Pewdiepie, was recruited to publish a promotional video, resulting in over 2 million views in 5 days. This advertising led to Sea Hero Quest becoming the most downloaded app on the Apple App store for a day and creating a 'moment' on Twitter. This snowball effect is what we attributed to being able to recruit 3.9 million participants drawn from every country in the world.

World-wide Participants

Since this dataset was recorded in a less controlled way than in a classic lab-based experiment, the first challenge was to filter out less dedicated players, for instance the ones who played a couple of levels with their friends around a beer and quit the game early. To ensure a good tradeoff between sample size and amount of data per player, we included in the analysis participants who played at least the first 6 wayfinding levels (level numbers 1, 2, 3, 6, 7 and 8) from a total of 45 levels and the first 2 path integration levels (levels 4 and 9). We only included participants who entered their demographics (57% of the total). Looking at the participant age distribution, we noticed two big spikes, at 18 and 99 years old. These were likely due to participants younger than 18 (whose data we couldn't collect for ethical reasons) or participants who didn't bother honestly entering their age, and instead chose one of the extremum of the spectrum. We discarded these participants. To further reduce selection bias and ensure stable cross-country comparisons, we only included participants from countries with at least 500 valid participants. For our first publication (Coutrot et al., 2018)(we kept collecting data for another year), this left us with 558,143 participants from 57 countries, 312,886 males (age: 34.97 ± 14.39 years old) and 245,257 females (age: 35.98 ± 15.50 years old).

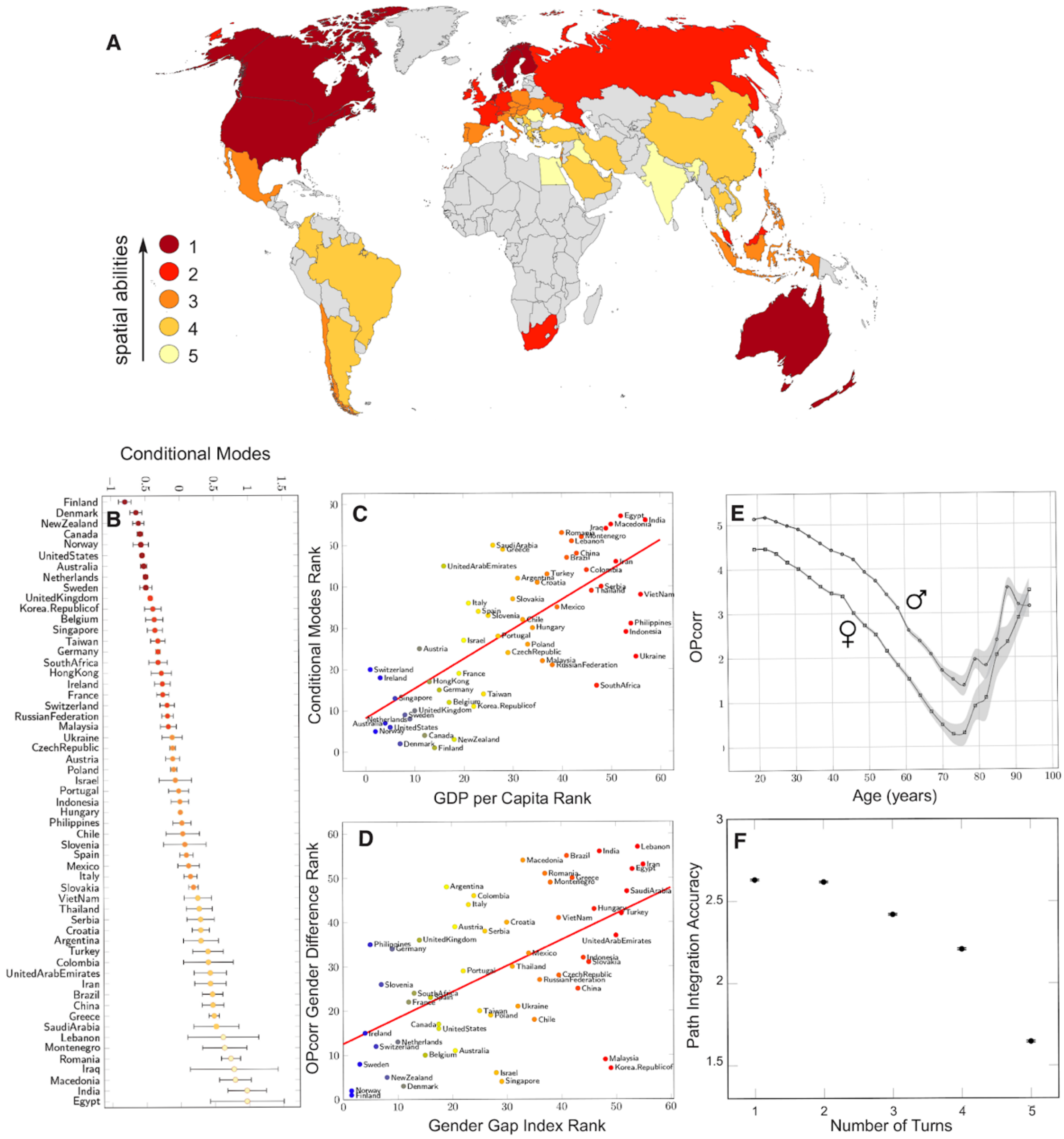

Figure 2 | Spatial Ability Distribution across Age, Gender, and Nations. a-b Five world clusters of people with similar Spatial Ability corrected for video gaming skill (SA). We used a multilevel model to predict SA with fixed effect for age and gender and random effect for nationality. Conditional modes (CMs) represent the country-level performance (the lower the better). **c -** Correlation between country performance (CM) and GDP per capita ($r = 0.69$, $p < 0.001$). **d -** Correlation between gender estimates and Gender Gap Index ($r = 0.62$, $p < 0.001$). **e -** Evolution of SA across age and gender. Data points correspond to the average OPcorr (Overall Performance, corrected by video games skill) within 3-year windows. Error bars correspond to SEs. **f -** Path-integration accuracy (number of stars) versus path complexity (number of turns). This plot includes participants that completed all five levels (N = 19,038). Error bars correspond to SEs. Reproduced from (Coutrot et al., 2018).

Quantifying spatial ability

Optimal navigation will travel the shortest distance to the goal (shortest length of the trajectory). Whereas inaccurate navigation will create long paths to the goal (Long trajectories). Thus, akin to many other studies (e.g. Spiers et al., 2001) we used the length of the trajectory as simple metric to quantify spatial ability. We considered that video game experience might bias performance, with players familiar with gaming on smartphones and tablets having an advantage in keeping the boat on the path chosen, rather than colliding with obstacles. Therefore, we normalized durations and trajectory lengths by dividing them by the sum of their values at the first two levels, where there are no choices and no need for spatial navigation. For more details on controlling familiarity with technology, see Coutrot et al (2018).

Effect of age, gender and nationality on spatial ability

The angles to start digging into this massive dataset - one of the biggest behavioral dataset recorded for scientific purposes - were almost endless. We decided to start with the most obvious ones, i.e. to check whether we could replicate some well-known results from the literature. We first verified whether spatial ability declined with age (Anguera et al., 2013; Ghisletta et al., 2012; Lindenberger, 2014), and whether males performed better than females (Linn and Petersen, 1985; Nazareth et al., 2019; Reilly and Neumann, 2013)). We fit a mixed model for spatial ability (SA), with fixed effects for age and varying slope for gender, nested within nationality: SA ~ age + (gender | country).

For age, across all included countries, we observed a similar pattern of a linear decline in spatial ability with age between 19 and 70 years old. However, after 70 years of age, performance counter-intuitively started to rise again. This was probably due to a strong selection bias in older participants: people above 70 y.o. playing video games for scientific research on their smartphones or tablets are likely in the upper band of their age group, cognitively speaking. This should be considered as a warning when generalizing effects measured in older participants to the general older population. This holds true in classic lab-based studies, which could lead to an even stronger selection bias, as physically going to a research facility to be tested can arguably constitute a stronger filter than simply playing games on a mobile device.

For gender, we observed an overall advantage of males over females (Cohen's d = 0.29, 95% CI = [0.27, 0.31]), consistent with the literature (Nazareth et al., 2019). However, this gender effect greatly varied across countries, from d = 0.09 in Norway to d = 0.48 in Argentina. These variations in effect size across countries can be associated with gender inequalities in society, assessed by the World Economic Forum's Gender Gap Index (GGI), which reflects economic and political opportunities, education, and well-being for women. We reported a positive correlation between GGI and the country-level gender effect size ($r = 0.62$, $p < 0.001$). This association remained significant after controlling for countries' GDP per capita, suggesting that the gender effect is not simply related to countries' wealth, but also to the improvement of the role of women in society. While biological differences between sexes may underly some of the gender differences in navigation performance, the weak effects sizes we found in several countries (e.g. Norway) suggest that societal factors such as access to education have a significant impact on gender differences in spatial navigation. More research to explore the impact of education on navigation ability and how this may be related to gender differences across nations is particularly warranted.

To assess the impact of nationality on spatial navigation, we fit a mixed model for Spatial Ability, with fixed effects for age and gender and random effect for nationality: SA ~ age + gender + (1 | country). The variance partition coefficient indicates that 1.7% of the variance in Spatial Ability can be attributed to differences between nationalities, after correcting for age and gender. We found that the country-level spatial ability is not geographically random, but strongly correlated with GDP per capita ($r = 0.69$, $p < 0.001$). This is likely due to numerous variables correlated with GDP per capita and associated with better spatial abilities, such as level of education (Skirbekk and Loichinger, 2012) - particularly in science (Gunderson et al., 2012; Uttal et al., 2013) - or ability to travel (Poumanyvong et al., 2012). However, country-level attributes other than GDP per capita and GGI can explain the observed distribution of spatial ability. For instance, the car culture in North America and Australia can explain why these countries are more successful than GDP equivalent European countries that rely more on public transport (Maguire et al., 2006; Sandamas and Foreman, 2015). However, this hypothesis fails to explain why the Nordic countries are at the top of the nation ranking. One possibility might be that the Nordic countries share a culture of participating in a sport related to navigation: orienteering. They have a tradition of teaching orienteering in schools (Annerstedt, 2008) and winning medals in the Orienteering World Championship. Across the 19 nations represented in the top 100 rankings, performance in orienteering is significantly correlated with countries' Conditional Modes (corrected country performance from our model) in SHQ (Pearson's correlation $r = 0.55$, $p = 0.01$), even after correcting for GDP per capita. This is a post-hoc hypothesis and should be considered with care, but we believe that this observation illustrates the potential for future targeted research to evaluate the impact of cultural activities on cognitive performance.

Does growing up in a city or outside cities impact navigation ability?

In addition to age, gender and nationality we have explored the impact of other demographics. One such demographic is whether participants reported growing up in a city or outside a city. We found that on average, a rural or mixed city/rural upbringing leads to better navigation ability than city upbringing, see Figure 3a (Coutrot et al., 2020). Notably, this result held when controlling for educational attainment. However, the negative effect of growing up in cities was mitigated by the topology of their street network (Figure 3b). The largest effect was observed in cities with a low Street Network Entropy (i.e. a regular layout, like Chicago, USA, see Boeing, 2019). In cities with high Street Network Entropy (i.e. more organic, like most old European city centres), this effect was much reduced. Moreover, we observed that participants who grew-up in cities with low Street Network Entropy performed better at less entropic SHQ levels. Altogether, these results suggest that people are better at navigating in environments topologically similar to where they grew up in, which highlights the importance of urban design on human cognition.

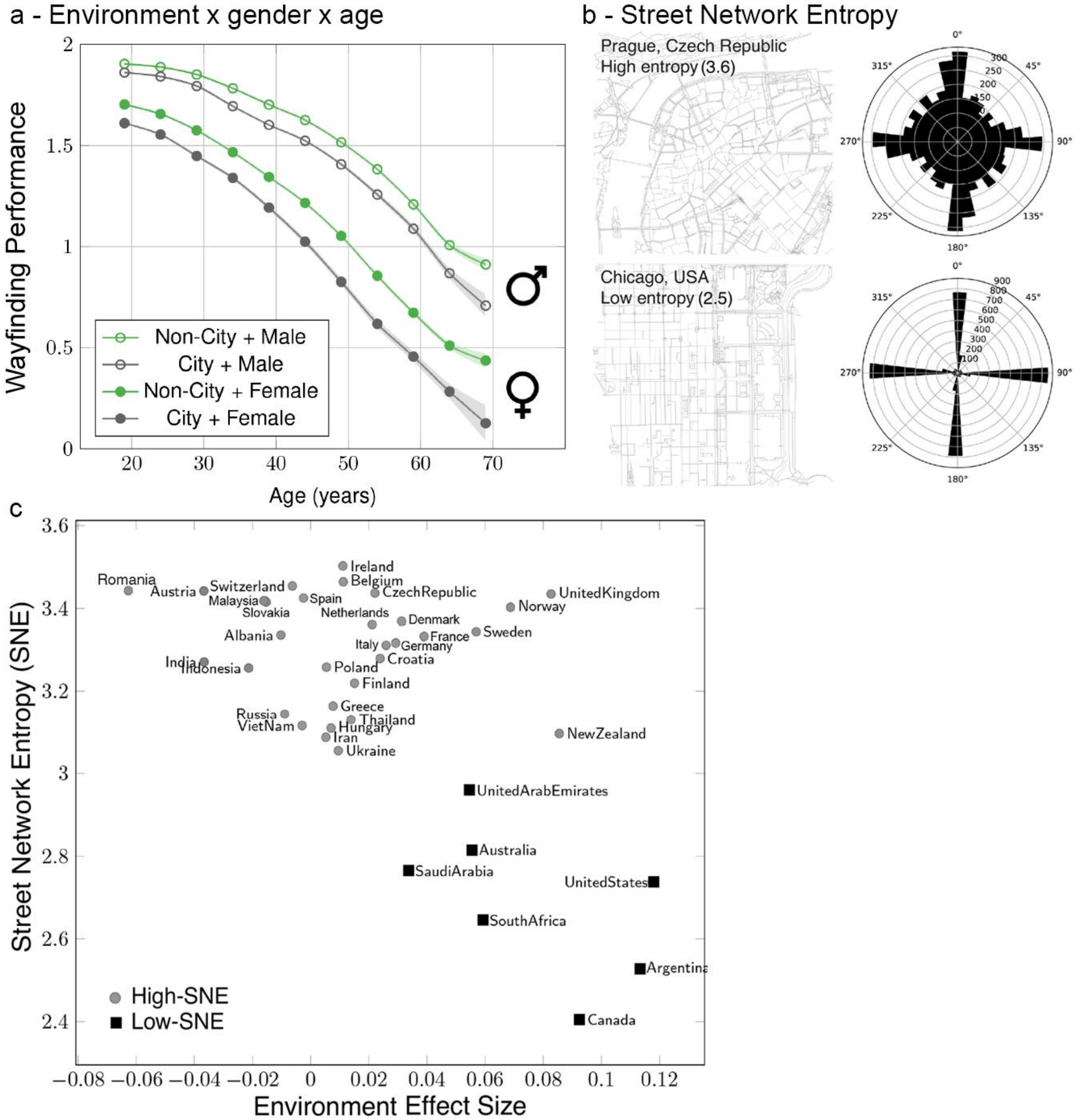

Figure 3 | Effect of the upbringing environment on spatial ability. a- Association between Environment and Wayfinding Performance stratified by age, gender, and education. The Wayfinding Performance is computed from the trajectory length and has been averaged within 5-years windows. Error bars represent standard error. **b-** Two examples cities with low (Chicago, USA) and high (Prague, Czech Republic) Street Network Entropy (SNE) (Boeing, 2019). **c-** Average SNE as a function of the environment effect size (random environment slope) in 38 countries. Positive values indicate an advantage of participants raised outside cities compared to their urban compatriots. Average SNE is the weighted average over the 10 most populated cities of the country, weighted by their population. Adapted from (Coutrot et al., 2020).

Real-world validity

Despite the effort we invested into designing a user-friendly game and analytically separating video gaming skills (speed of tutorial completion) from spatial ability, we considered that it was plausible that Sea Hero Quest does not accurately reflect real-world navigation skill. To navigate in Sea Hero Quest you need to update our perspective in the game coordinates rather than with whole-body rotations in the real-world. Notably, previous studies found that navigation performance in the real world correlates with navigation performance in virtual environments (Conroy, 2001; Cushman et al., 2008; Moffat et al., 2001), they mainly used desktop or immersive displays, while our dataset has been collected via a mobile video-game. To directly assess the generalizability of our data to the real world, we tested participants with Sea Hero Quest and with a similar real-world navigation task in the streets of London (UK) and Paris (France) (Coutrot et al., 2019). We found a significant correlation between real-world navigation performance and performance on SHQ levels that specifically test navigation (in London $r = 0.46$, $p = 0.01$; in Paris $r = 0.57$, $p = 0.001$). No such correlation was found when comparing performance on the tutorial (where spatial ability was not required) levels with real-world navigation performance ($r = 0.06$). Thus, at least in these urban settings, Sea Hero Quest seems to capture real-world navigation performance.

Understanding how the environment impacts navigation performance

The environment can have a dramatic impact on how difficult it is to navigate. Walking to the next room is easier than a trek across the Sahara desert. A core research question is what features of the environment are particularly important in determining difficulty. Key factors appear to be: the size of the environment, the number of decision points, the number of regions of space, the presence of distinct fixed landmarks, and the organisation/disorganisation of the space (Arthur and Passini, 1992; Raubal and Egenhofer, 1998; Conroy, Hoelscher & Spiers 2015; Ekstrom et al., 2018).

The majority of Sea Hero Quest wayfinding levels were designed by Ruth Dalton and Christoph Hoelscher to vary in these properties, allowing analysis of how wayfinding behaviour is related to such factors (Coutrot et al. 2018). Our preliminary analysis indicates that indeed the number of decision points and topological structure of the environment has an impact on wayfinding difficulty (Yesiltepe et al., 2019a). Global landmarks have less of an impact on navigation than we had predicted (Yesiltepe et al., 2019b). Exploring this further it was possible to quantify the distinction between local and global landmarks by the extent to which they were in view during the navigation task (Yesiltepe et al., 2020a). Furthermore, we found that seemingly insignificant features (such as a cluster of reeds) could be considered important landmarks and that current models of extracting saliency in images were unable to predict what would be considered a useful landmark (Yesiltepe et al., 2020b,c,d). This highlights the importance of location (e.g. at waypoints) of landmarks for navigation (Caduff & Timpf, 2008). Future research in this direction will explore how individual differences interact with the environment features to explain navigational choices and trajectory patterns.

Is Sea Hero Quest reliable for test-retest?

A key aspect in the development of new cognitive tests, is their reliability not only to detect the effects of interest but also whether those effects can be elicited repeatedly across test sessions. Such test-retest reliability is commonly established for clinical tests which require longitudinal follow-up of patients, however, it is rarely conducted for experimental cognitive tests. Since SHQ has the potential to track people's navigation performance over time and to potentially use this as a clinical marker for incipient Alzheimer's disease, establishment of test-retest reliability was a key element for our SHQ research. Specially to determine the degree to which baseline and follow-up assessments produce consistent results. For example, participants may learn to apply strategies at retest that improve their navigation accuracy or efficiency, reducing its test-retest reliability.

We therefore conducted a test-retest study with middle-aged people, who underwent SHQ testing at baseline and an average follow-up at 18 months (Coughlan et al., 2020a). The retest delay was chosen to reflect appropriate time windows to follow up people for long-term longitudinal studies. The results revealed that some measures of SHQ showed moderate test-retest reliability, such as distance travelled, whereas others showed only low test-retest reliability, such as SHQ duration. Interestingly, this reliability could also be confirmed for 'at-genetic-risk' of Alzheimer's disease participants, who showed similar results.

Overall, SHQ shows good test-retest reliability, similar to existing spatial navigation tests such as the 4 Mountains Test (Chan et al., 2016). However, the choice of outcome variable impacts the reliability measurements, with distance travelled emerging as the most reliable measure to track navigation changes in SHQ over time.

Can SHQ be helpful to monitor variation in skill for those at genetic risk of AD?

The test-re-test findings also highlight that at-genetic-risk of AD participants can be successfully tracked over time, which has clear implications for incipient disease monitoring. To investigate the at-genetic-risk angle further, we conducted a study in a large cohort of participants, either at genetic risk or not (Coughlan et al., 2019).

The genetic risk factor we selected as the most critical was APOE (apolipoprotein E). APOE is the most common genetic risk factor for AD, with epidemiological studies showing that carrying the heterozygous form (e3:e4) increases people's risk 4-fold for AD and the homozygous form (e4:e4) increases the risk 12-fold, compared to the wild type (e3:e3) APOE alleles (Corder et al., 1993). The exact reasons why APOE increases risk for Alzheimer's disease are still being investigated, but it has been implicated with potential changes to blood-brain barrier integrity and reduced microglia activity, thus slowing down the removal of beta-amyloid fibrils - a key culprit in the onset of Alzheimer's disease.

There has been an increasing interest in the APOE cohort as a disease model for incipient cognitive changes for Alzheimer's disease. Previous studies have shown that APOE carriers at increased genetic risk show changed navigation behaviour with overshooting of navigation targets and preference for border areas (Coughlan et al., 2020b). It has been speculated that APOE induced grid cell changes in the entorhinal cortex are responsible for these specific navigation changes, as distance travelled and landmark spatial re-setting are critical factors mediated by the entorhinal cortex (Kunz et al., 2015).

We replicated the previous finding with SHQ, by showing that at-genetic-risk (e3:e4 & e4:e4) participants travelled longer distances within the game than APOE wildtype (e3:e3) participants (Coughlan et al., 2019). Further, we replicated the finding that at-genetic-risk participants travelled closer to the borders within SHQ, potentially to compensate for incipient grid cell linked changes in entorhinal cortex. A critical advantage which SHQ has to previous studies is that the same APOE navigation effect can be detected with virtually no training and much quicker. Previous experimental designs relied on participants learning spatial environments over repeated trials, with some learning paradigms taking as long as 1 hour. By contrast, SHQ requires no learning of environments, but still results in the same effects, allowing inclusion with other cognitive tests for dementia assessment (Coughlan et al., 2018b). This is also consistent with the premise that the cognitive problems linked to the APOE gene variants are related to spatial disorientation rather than learning per se.

The other final advantage which SHQ has, is that it allows for a more personalised detection of navigation changes. Not only can we detect group differences between genetic-risk groups, but we can also compare each at-genetic-risk participant to their age, gender and country matched peers, via the large normative SHQ data set. Even for a single age, gender and country, the large data of SHQ allows to compare some at-risk of AD against at least 700 healthy peers. What additional information does this individual comparison provide? The key additional information is that it allows to benchmark at-risk of AD people against their peers, thus showing whether they perform within the benchmark navigation performance or further off it. The rationale being that the further off people at-risk of AD perform, the more likely their navigation specific brain regions are affected by incipient AD pathophysiology. The distance from the mean performance allows, therefore, to potentially determine how close people are to being at-risk of developing Alzheimer's disease, which has clear diagnostic, prevention and treatment implications.

Using machine learning to learn efficient representations of the trajectories

So far, we quantified the spatial ability measured in Sea Hero Quest with very simple metrics, mainly based on the length of the trajectories followed by the participants to complete the levels. The longer the trajectory, the more participants got lost, the worse the performance. Although very intuitive, this metric fails at capturing many aspects of spatial ability. For instance, two participants could follow two very different trajectories with exactly the same length. Our metric would not be able to separate these participants, even though they used completely different navigation strategies. To overcome this issue, we need to develop more sophisticated approaches to represent and quantify the strategies developed by the players. Spatial trajectories are spatial patterns that contain a wealth of information arising from the complex interaction between the environment and the navigator's decisions. Several trajectory-mining approaches have been proposed in many fields such as traffic forecasting, object motion prediction, ecology, behavioural and clinical neuroscience (Damascène Mazimpaka and Timpf, 2016). The most popular approach is trajectory clustering, which aims at discovering groups of similar trajectories in terms of intuitive features such as their coordinates, angle, speed, or curvature (Bian et al., 2018). Another approach is to learn efficient representations of the trajectories directly from the raw data, using Artificial Neural Networks (ANNs) (Wang et al., 2019). ANNs can learn more efficient representations than hand-crafted features, but they are often black-boxes and it is hard to interpret the concepts learnt by the models. We recently proposed an ANN-based model learning trajectory representations and connecting them to participant's demographics in an interpretable way (Dubois et al., 2021). It relies on three parallel modules capturing different aspects of the signal (spatial, temporal, and spectral), allowing the user to understand which parts of the trajectories most contribute to the output. Future studies will use these approaches to further understand the origins of the differences in performance across the demographics.

Addressing the limitations and challenges of using mobile apps for research

Using a mobile app to examine navigation ability has provided an unprecedented capacity to understand human behaviour on a mass scale. However, there are several limitations and challenges with this approach. Some of the key challenges are: the real-world validity of a mobile phone test, bias in the sample of participants, the limited demographic questions posed, and the trade off between a highly engaging video game and rigorous cognitive testing. We have discussed previously our exploration of real-world validity of our test, below we consider the other issues.

All human psychology studies have a bias in who volunteers to take part (Henrich, Heine & Norenzayan, 2010). Mobile apps combined with a citizen science provide the capacity to reach out beyond university students, broader age ranges, people with less education and across different countries. However, people in older age groups will be less likely to engage in downloading video games to phones and tablet devices, thus skewing the sample to a younger age. Fortunately for Sea Hero Quest, the mass popularity meant that even with this skew we were able to sample 100,000s of people over 50 years and 1000s over 80 years across many countries and from people with lower educational backgrounds (Coutrot et al., 2018). However, as discussed previously we observed a significant selection bias after the age of around 70 years (Fig 2E). After the age of 70 years, performance steadily increases with age. Given the wealth of evidence for a negative impact of aging on cognition (Lindenberger 2014) this result is not plausible as a direct measure of the general populace. Rather it seems more likely that post 70 participants who are willing to engage in a research based app game will increasingly be healthy agers who are motivated to participate in studies. See Coutrot et al (2018) supplemental information for details on our analysis of selection bias in the data. Documenting this bias is important as we believe that if it is present in our data, it is also likely present in conventional lab-based experiments. Indeed, the filter represented by playing a game on a smartphone might not be stronger than the one represented by asking participants to physically come to a research facility.

Unlike some other apps aimed for research, such *The Great Brain Experiment (*Smittenaar et al, 2015), Sea Hero Quest was designed to be an engaging video game that people would play purely for fun whilst also generating data. This was done in order to maximise the sample size from the data collected. As part of this design, rather than halt game play with many demographics questions, we focused on key demographics and the players were awarded ‘virtual badges’ for answering the questions. The questions were also framed with images to make them more comprehensible and to engage participants with the questions. We ideally would have invited participants to provide answers to more questions via an online survey. However, because Sea Hero Quest was designed to securely store anonymised data directly from the app we were not able at the time of launching the app to find a way to link participants to a web-based survey of further demographics. Future research will be able to address this.

An important limitation from this research is to consider that Sea Hero Quest examines a particular set of spatial skills, such as interpreting a map and memorising it to navigate to a set of goals in the right order. People in traditional communities can navigate with exceptional skill in highly complex environments such as oceans, deserts and tundra (Levinson, 2003; Ekstrom et al., 2018). Such environments are learned from early life and are encountered over many years of experience. Individual variation in such skill is not something an app on a mobile device can easily aid our understanding of. It will be important in future work to explore the links between the skill of highly expert navigators such as London taxi drivers (Maguire et al. 2006; Spiers & Maguire 2008; Greisbauer et al., 2021). Moreover, it would be beneficial to understand how navigation performance in Sea Hero Quest relates to other spatial and non-spatial performance, given evidence that spatial navigation may be distinct from performance on more object-based spatial tasks (Malanchini et al., 2020), non-spatial association memory (Ngo et al., 2016) or autobiographical memory (Fan et al., 2021). Our large online sample from Sea Hero Quest can separate spatial navigation ability from video games skill (tutorial levels), but we cannot dissociate it from general cognitive ability, 'g', as it was not possible to incorporate other general IQ tasks with our funded time-lines for the research. However, past research suggests a large portion of the variance in spatial navigation is separable from measures of general cognitive ability (Malanchini et al., 2020). Thus, we would expect our measure of spatial navigation ability to be separable to a significant extent from a general cognitive ability.

Future Directions

There are a range of different future advances that may come from this approach to understanding navigation. It will be important to understand how the other demographics we collected impact navigation skills. These include the impact of education, handedness, sleep and self-rated navigation ability. Future research exploring machine learning and approaches to trajectory prediction hold much promise for examining the large trajectory data created by Sea Hero Quest. This research might for example identify that participants become less predictable with age. It could be that men are more predictable in Sea Hero Quest environments lacking landmarks (Reilly & Neumann 2013). Data collected from an immersive VR version of Sea Hero Quest for the Samsung Gear VR device has yet to be examined in detail and may shed more light on the disorientation in relation to various individual differences.

Combining Sea Hero Quest with neuroimaging or neural recording methods will be useful for exploring whether brain structure or function can predict performance metrics from Sea Hero Quest. Current research is divided on whether variation in hippocampal structural measures can be linked to performance; with some studies reporting a correlation (Bohbot et al., 2007; Schinazi et al., 2013; Hartley & Harlow, 2014; Brunec et al., 2019; Hodgetts et al., 2020) and others reporting no correlation (Maguire et al., 2003; Weisberg, Newcombe & Chaterjee, 2019; Clark et al., 2020). Sea Hero Quest would be beneficial in that the participants can be selected to cover an useful range in distribution of individual differences. Past studies may have studied too few individuals with poor navigation skills or high navigation skills making it difficult to explore the full range. Functional imaging studies have shown important individual differences in activity patterns (Hartley et al., 2003; Wolbers and Hegarty, 2010; Ekstrom et al., 2018; Patai et al., 2019). Sea Hero Quest would allow quantification of a range of variables linked both to the variation in environments used in the game (Yesiltepe et al., 2019a) as well as other graph-theoretic metrics found to be related to hippocampal activity, such as degree centrality (Javadi et al., 2017). Combining neuroimaging data with the mass trajectory data may allow exploration of whether patterns of activity in key brain regions enhance forecasting of future trajectories, potentially predicting when individuals will back-track (Javadi et al, 2019). Similarly, recording the gaze behaviour of participants playing Sea Hero Quest would be particularly interesting since spatial information acquisition happens in large part through the visual sense.

One missing line of research so far in spatial cognition has been the use of longitudinal cohort studies. Testing thousands of such individuals would provide links to much broader and detailed assessments of performance over a life-time and allow for analysis of changes in performance as these cohorts are followed up. So far Sea Hero Quest has explored spatial navigation over the age of 18 and we have found a general decline in performance. This raises the question: at what age does navigation ability peak? Hopefully future studies with those under 18 will help address this key question. At the other end of life we hope that Sea Hero Quest will be a useful tool for monitoring cognitive function in Alzheimer's disease (Couglan et al., 2020b).

Conclusions

The global reach of mobile phone based game applications has enabled the mass testing of millions of people in a manner not feasible a decade ago (Gillan & Rutledge, 2021). Our project Sea Hero Quest has shown that it is possible to assess the impact of individual differences in spatial navigation ability across 3.9 million people, sampling population data in 57 countries (Coutrot et al., 2018; Coutrot et al., 2020). Our results replicated prior small sample studies: age had a strong negative impact on performance, there was an advantage for males and those with greater education (Coutrot et al., 2018; Coutrot et al., 2020). Extending beyond prior studies with population data we were able to show that such patterns also exist in non-western countries, and that the economic wealth of a country (GDP) predicts its population's average navigation performance (Coutrot et al., 2018). We were also able to reveal a relationship between gender inequality of a country and how large the gender disparity was in the spatial navigation skill (Coutrot et al. 2018), indicating that cultural differences in how men and women are treated impact spatial navigation ability. Most recently we have revealed that growing up in the city of countries with griddy cities (e.g. Chicago) has an overall negative impact on navigation skill in the game (Coutrot et al. 2020). We have also briefly discussed how Sea Hero Quest has been useful for monitoring changes in skill for those at risk of Alzheimer's disease (Coughlan et al., 2019; 2020b) and exploring broader metrics of spatial performance with artificial neural networks (Dubois et al., 2021). Future research with Sea Hero Quest will allow a variety of studies to be done remotely with healthy and clinical populations. Analysis of this rich data has still much to reveal in helping address the individual differences in spatial navigation ability.

Acknowledgements

We are grateful to various collaborators for their input to the project: Ed Manley, Jan Wiener, Ruth Conroy Dalton, Christoph Hölscher, Véronique Bohbot, Ricardo Silva, Gillian Coughlan, Zita Patai, Will de Cothi, Hippolyte Dubois. We thank Lara Gregorians for comments on a draft. We thank Deutsche Telekom for supporting and funding this research, Glitchers Ltd for game production, Saatchi and Saatchi London for project management and creative input, Alzheimer’s Research UK for funding the analysis (ARUK-DT2016-1), and Tim Parry at ARUK for supporting us through all the challenges.

References

Alzheimer’s Disease International (2018). World Alzheimer report. The state of the art of dementia research: New frontiers.

Anguera, J. A., Boccanfuso, J., Rintoul, J. L., Al-Hashimi, O., Faraji, F., Janowich, J., Kong, E., Larraburo, Y., Rolle, C., Johnston, E., and Gazzaley, A. (2013). Video game training enhances cognitive control in older adults. Nature, 501:97–101.

Annerstedt, C. (2008). Physical education in scandinavia with a focus on sweden: a comparative perspective. Physical Education and Sport Pedagogy, 13(4):303–318.

Arthur, P., & Passini, R. (1992). *Wayfinding: people, signs, and architecture*. McGraw-Hill.

Bian, J., Tian, D., Tang, Y., and Tao, D. (2018). A survey on trajectory clustering analysis. arXiv:1802.06971, pages 1–40.

Boeing, G. (2019). Urban spatial order: Street network orientation, configuration, and entropy. *Applied Network Science*, *4*(1), 1-19.

Bohbot, V. D., Lerch, J., Thorndycraft, B., Iaria, G., & Zijdenbos, A. P. (2007). Gray matter differences correlate with spontaneous strategies in a human virtual navigation task. *Journal of Neuroscience*, *27*(38), 10078-10083.

Brunec, I. K., Robin, J., Patai, E. Z., Ozubko, J. D., Javadi, A. H., Barense, M. D., ... & Moscovitch, M. (2019). Cognitive mapping style relates to posterior–anterior hippocampal volume ratio. *Hippocampus*, *29*(8), 748-754.

Burles, F. and Iaria, G. (2020). Behavioural and cognitive mechanisms of developmental topo-graphical disorientation. Scientific reports, 10(1):1–11.

Caduff, D., & Timpf, S. (2008). On the assessment of landmark salience for human navigation. *Cognitive processing*, *9*(4), 249-267.

Chan, D., Gallaher, L. M., Moodley, K., Minati, L., Burgess, N., & Hartley, T. (2016). The 4 mountains test: a short test of spatial memory with high sensitivity for the diagnosis of pre-dementia Alzheimer's disease. JoVE (Journal of Visualized Experiments), 116:e54454.

Clark, I. A., Monk, A. M., Hotchin, V., Pizzamiglio, G., Liefgreen, A., Callaghan, M. F., & Maguire, E. A. (2020). Does hippocampal volume explain performance differences on hippocampal-dependant tasks?. *NeuroImage*, *221*, 117211.

Conroy, R. A. (2001). Spatial navigation in immersive virtual environments. PhD thesis, University College London.

Dalton, R. C., Hoelscher, C., & Spiers, H. J. (2015). Navigating complex buildings: Cognition, neuroscience and architectural design. In *Studying visual and spatial reasoning for design creativity* (pp. 3-22). Springer, Dordrecht.

Corder, E. H., Saunders, A. M., Strittmatter, W. J., Schmechel, D. E., Gaskell, P. C., Small, G., Roses, A., Haines, J., and Pericak-Vance, M. A. (1993). Gene dose of apolipoprotein e type 4 allele and the risk of Alzheimer’s disease in late onset families. Science, 261(5123):921–923.

Coughlan, G., Laczo, J., Hort, J., Minihane, A.-M., and Hornberger, M. (2018a). Spatial navigation deficits — overlooked cognitive marker for preclinical Alzheimer disease? Nature Reviews Neurology, 14:496–506.

Coughlan, G., Flanagan, E., Jeffs, S., Bertoux, M., Spiers, H., Mioshi, E., & Hornberger, M. (2018b). Diagnostic relevance of spatial orientation for vascular dementia: A case study. *Dementia & neuropsychologia*, *12*(1), 85-91.

Coughlan, G., Coutrot, A., Khondoker, M., Minihane, A.-M., Spiers, H., and Hornberger, M. (2019). Toward personalized cognitive diagnostics of at-genetic-risk Alzheimer's disease. Proceedings of the National Academy of Sciences, 116(19):9285–9292.

Coughlan, G., Puthusseryppady, V., Lowry, E., Gillings, R., Spiers, H. J., Minihane, A.-M., and Hornberger, M. (2020a). Test-retest reliability of spatial navigation in at-risk Alzheimer's disease. PLoS ONE, 15(9).

Coughlan, G., Zhukovsky, P., Puthusseryppady, V., Gillings, R., Minihane, A. M., Cameron, D., & Hornberger, M. (2020b). Functional connectivity between the entorhinal and posterior cingulate cortices underpins navigation discrepancies in at-risk Alzheimer's disease. *Neurobiology of aging*, *90*, 110-118.

Coutrot, A., Silva, R., Manley, E., de Cothi, W., Sami, S., Bohbot, V. D., Wiener, J. M., Hölscher, C., Dalton, R. C., Hornberger, M., and Spiers, H. J. (2018). Global Determinants of Navigation Ability. Current Biology, 28(17):2861–2866.

Coutrot, A., Schmidt, S., Coutrot, L., Pittman, J., Hong, L., Wiener, J. M., Hölscher, C., Dalton, R. C., Hornberger, M., and Spiers, H. J. (2019). Virtual navigation tested on a mobile app is predictive of real-world wayfinding navigation performance. PLoS ONE, 14(3):1–15.

Coutrot, A., Manley, E., Yesiltepe, D., Dalton, R. C., Wiener, J. M., Hölscher, C., Hornberger, M., and Spiers, H. J. (2020). Cities have a negative impact on navigation ability: evidence from 38 countries. bioRxiv, https://www.biorxiv.org/content/10.1101/2020.01.23.917211v3.

Cushman, L. A., Stein, K., and Duffy, C. J. (2008). Detecting navigational deficits in cognitive aging and Alzheimer disease using virtual reality. Neurology, 71(12):888–895.

Damascène Mazimpaka, J. and Timpf, S. (2016). Trajectory data mining: A review of methods and applications. Journal of Spatial Information Science, (13):61–99.

Dubois, H., Le Callet, P., Hornberger, M., Spiers, H. J., and Coutrot, A. (2021). Capturing and Explaining Trajectory Singularities using Composite Signal Neural Networks. In European Conference on Signal Processing (EUSIPCO).

Ekstrom, A. D., Spiers, H. J., Bohbot, V. D., and Rosenbaum, R. S. (2018). Human spatial navigation. Princeton University Press.

Fan, C. L., Abdi, H., & Levine, B. (2021). On the relationship between trait autobiographical episodic memory and spatial navigation. *Memory & Cognition*, *49*(2), 265-275.

Ghisletta, P., Rabbitt, P., Lunn, M., and Lindenberger, U. (2012). Two thirds of the age-based changes in fluid and crystallized intelligence, perceptual speed, and memory in adulthood are shared. Intelligence, 40:260–268.

Gillan, C. M., & Rutledge, R. B. (2021). Smartphones and the neuroscience of mental health. *Annual Review of Neuroscience*, *44*.

Griesbauer, E. M., Manley, E., Wiener, J. M., & Spiers, H. J. (2021). Learning The Knowledge: How London Taxi Drivers Build Their Cognitive Map of London. *bioRxiv*.

Gunderson, E. A., Ramirez, G., Beilock, S. L., and Levine, S. C. (2012). The relation between spatial skill and early number knowledge: the role of the linear number line. Developmental Psychology, 48(5):1229.

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*(5), 877-888.

Hartley, T., & Harlow, R. (2012). An association between human hippocampal volume and topographical memory in healthy young adults. *Frontiers in human neuroscience*, *6*, 338.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world?. Behavioral and brain sciences, 33(2-3), 61-83.

Hodgetts, C. J., Stefani, M., Williams, A. N., Kolarik, B. S., Yonelinas, A. P., Ekstrom, A. D., ... & Graham, K. S. (2020). The role of the fornix in human navigational learning. *cortex*, *124*, 97-110.

Javadi, A. H., Emo, B., Howard, L. R., Zisch, F. E., Yu, Y., Knight, R., ... & Spiers, H. J. (2017). Hippocampal and prefrontal processing of network topology to simulate the future. *Nature communications*, *8*(1), 1-11.

Javadi, A. H., Patai, E. Z., Marin-Garcia, E., Margois, A., Tan, H. R. M., Kumaran, D., ... & Spiers, H. J. (2019). Backtracking during navigation is correlated with enhanced anterior cingulate activity and suppression of alpha oscillations and the ‘default-mode’ network. *Proceedings of the Royal Society B*, *286:* 20191016.

Kunz, L., Schröder, T. N., Lee, H., Montag, C., Lachmann, B., Sariyska, R., Reuter, M., Stirnberg, R., Stöcker, T., Messing-Floeter, P. C., Fell, J., Doeller, C. F., and Axmacher, N. (2015). Reduced grid-cell–like representations in adults at genetic risk for Alzheimer’s disease. Science, 350(6259):430–433.

Laczo, J., Andel, R., Vyhnalek, M., Vlcek, K., Nedelska, Z., Matoska, V., Gazova, I., Mokrisova, I., Sheardova, K., and Hort, J. (2014). Apoe and spatial navigation in amnestic MCI: Results from a computer-based test. Neuropsychology, 28(5):676.

Levinson, S. C. (2003). *Space in language and cognition: Explorations in cognitive diversity* (No. 5). Cambridge University Press.

Lindenberger, U. (2014). Human cognitive aging: Corriger la fortune? Science, 346(6209):572–578.

Linn, M. C. and Petersen, A. C. (1985). Emergence and characterization of sex differences in spatial ability: A meta-analysis. Child Development, 56(6):1479–1498.

Maguire, E. A., Spiers, H. J., Good, C. D., Hartley, T., Frackowiak, R. S., & Burgess, N. (2003). Navigation expertise and the human hippocampus: a structural brain imaging analysis. *Hippocampus*, *13*(2), 250-259.

Maguire, E. A., Woollett, K., and Spiers, H. J. (2006). London taxi drivers and bus drivers: a structural MRI and neuropsychological analysis. Hippocampus, 16(12):1091–1101.

Malanchini, M., Rimfeld, K., Shakeshaft, N. G., McMillan, A., Schofield, K. L., Rodic, M., ... & Plomin, R. (2020). Evidence for a unitary structure of spatial cognition beyond general intelligence. *npj Science of Learning*, *5*(1), 1-13.

Moffat, S. D., Zonderman, A. B., and Resnick, S. M. (2001). Age differences in spatial memory in a virtual environment navigation task. Neurobiology of Aging, 22:787–796.

Morgan, J. (2016). Gaming for dementia research: a quest to save the brain. *The Lancet Neurology*, *15*(13), 1313.

Morris, R. G., Garrud, P., Rawlins, J. a., and O'Keefe, J. (1982). Place navigation impaired in rats with hippocampal lesions. Nature, 297(5868):681–683.

Nazareth, A., Huang, X., Voyer, D., and Newcombe, N. (2019). A meta-analysis of sex differences in human navigation skills. Psychonomic bulletin & review, 26(5):1503–1528.

Ngo, C. T., Weisberg, S. M., Newcombe, N. S., & Olson, I. R. (2016). The relation between navigation strategy and associative memory: An individual differences approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(4), 663.

Patai, E. Z., Javadi, A. H., Ozubko, J. D., O'Callaghan, A., Ji, S., Robin, J., ... & Spiers, H. J. (2019). Hippocampal and retrosplenial goal distance coding after long-term consolidation of a real-world environment. *Cerebral Cortex*, *29*(6), 2748-2758.

Poumanyvong, P., Kaneko, S., and Dhakal, S. (2012). Impacts of urbanization on national transport and road energy use: Evidence from low, middle and high income countries. Energy Policy, 46:268–277.

Raubal, M., & Egenhofer, M. J. (1998). Comparing the complexity of wayfinding tasks in built environments. *Environment and Planning B: Planning and Design*, *25*(6), 895-913.

Reilly, D. and Neumann, D. L. (2013). Gender-role differences in spatial ability: A meta-analytic review. Sex Roles, 68(9-10):521–535.

Sandamas, G. and Foreman, N. (2015). Active Versus Passive Acquisition of Spatial Knowledge While Controlling a Vehicle in a Virtual Urban Space in Drivers and Non-Drivers. SAGE Open, 5(3):1–9.

Schinazi, V. R., Nardi, D., Newcombe, N. S., Shipley, T. F., & Epstein, R. A. (2013). Hippocampal size predicts rapid learning of a cognitive map in humans. *Hippocampus*, *23*(6), 515-528.

Skirbekk, V. and Loichinger, E. (2012). Variation in cognitive functioning as a refined approach to comparing aging across countries. Proceedings of the National Academy of Sciences, 109(3).

Smittenaar, P., Rutledge, R. B., Zeidman, P., Adams, R. A., Brown, H., Lewis, G., & Dolan, R. J. (2015). Proactive and reactive response inhibition across the lifespan. *PLoS One*, *10*(10), e0140383.

Spiers, H. J., Burgess, N., Hartley, T., Vargha-Khadem, F., & O'Keefe, J. (2001). Bilateral hippocampal pathology impairs topographical and episodic memory but not visual pattern matching. *Hippocampus*, *11*(6), 715-725.

Spiers, H. J., & Maguire, E. A. (2008). The dynamic nature of cognition during wayfinding. *Journal of environmental psychology*, *28*(3), 232-249.

Tu, S., Spiers, H. J., Hodges, J. R., Piguet, O., & Hornberger, M. (2017). Egocentric versus allocentric spatial memory in behavioral variant frontotemporal dementia and Alzheimer’s disease. *Journal of Alzheimer's disease*, *59*(3), 883-892.

Uttal, D. H., Meadow, N. G., Hand, E. T. L. L., Warren, C., and Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. Psychological Bulletin, 139(2):352–402.

Wang, S., Cao, J., and Yu, P. S. (2019). Deep Learning for Spatio-Temporal Data Mining: A Survey. arXiv:1906.04928, pages 1–21.

Weisberg, S. M., Schinazi, V. R., Newcombe, N. S., Shipley, T. F., & Epstein, R. A. (2014). Variations in cognitive maps: understanding individual differences in navigation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *40*(3), 669.

Weisberg, S. M., & Newcombe, N. S. (2016). How do (some) people make a cognitive map? Routes, places, and working memory. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *42*(5), 768.

Weisberg, S. M., Newcombe, N. S., & Chatterjee, A. (2019). Everyday taxi drivers: Do better navigators have larger hippocampi?. *Cortex*, *115*, 280-293.

Wiener, J. M., Büchner, S. J., and Hölscher, C. (2009). Taxonomy of human wayfinding tasks: A knowledge-based approach. Spatial Cognition & Computation, 9(2):152–165.

Wolbers, T., & Hegarty, M. (2010). What determines our navigational abilities?. *Trends in cognitive sciences*, *14*(3), 138-146.

Yesiltepe, D., Dalton, R., Ozbil Torun, A., Dalton, N., Noble, S., Hornberger, M., & Spiers, H. (2019a). A Wayfinding Research in Virtual Environments: The effect of spatial structure and different conditions on movement. Proceedings of the12th Space Syntax Symposium.

Yesiltepe, D., Dalton, R., Ozbil, A., Dalton, N., Noble, S., Hornberger, M., Coutrot A., & Spiers, H. (2019b). Usage of Landmarks in Virtual Environments for Wayfinding: Research on the influence of global landmarks. Proceedings of the12th Space Syntax Symposium.

Yesiltepe, D., Dalton, R. C., Torun, A. O., Noble, S., Dalton, N., Hornberger, M., & Spiers, H. (2020a). Redefining Global and Local Landmarks: When Does a Landmark Stop Being Local and Become a Global One?. In *German Conference on Spatial Cognition* (pp. 111-121). Springer, Cham.

Yesiltepe, D., Ozbil Torun, A., Coutrot, A., Hornberger, M., Spiers, H., & Conroy Dalton, R. (2020b). Computer models of saliency alone fail to predict subjective visual attention to landmarks during observed navigation. *Spatial Cognition & Computation*, 1-28.

Yesiltepe, D., Dalton, R. C., Torun, A. O., Coutrot, A., Hornberger, M., & Spiers, H. (2020c). A Study on Visual and Structural Characteristics of Landmarks and Experts' and Non-experts' Evaluations. In *German Conference on Spatial Cognition* (pp. 95-107). Springer, Cham.

Yesiltepe, D., Dalton, R. C., Torun, A. O., Hornberger, M., & Spiers, H. (2020d). Understanding Cognitive Saliency by Using an Online Game. In *German Conference on Spatial Cognition* (pp. 76-87). Springer, Cham.