Acta Psychologica 199 (2019) 102895

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

Gender, videogames and navigation in virtual space☆

Suzanne de Castell[a,*], Hector Larios[a], Jennifer Jenson[b]

[a] University of Ontario Institute of Technology, 11 Simcoe Street N., L1H 7L7 Oshawa, Ontario, Canada

[b] York University, 4700 Keele St., M3J 1P3 Toronto, Ontario, Canada

ABSTRACT

Spatial abilities associated with success in educational and occupational fields of Science, Technology, Engineering, and Mathematics (STEM) have been repeatedly shown to be gendered, with males demonstrating measurably better spatial abilities than females. Less is known about why this is, or about how experience with spatial systems (videogames, for example) affects these abilities. We conducted two experiments with 82 participants with varying degrees of videogame experience on measures of mental rotation, spatial learning, and spatial memory. Spatial learning and memory were tested in a Virtual Morris Water Maze. In the first experiment, the maze lacked proximal landmarks. Males proved faster and more accurate than females in learning the location of the hidden platform. As predicted males also outperformed females in mental rotation abilities. Mental rotation correlated with performance in the virtual maze, indicating that in the absence of proximal landmarks, participants relied on strategies requiring mental rotation. Experienced 3D videogame players did not demonstrate superior spatial learning and memory, but performed better than novices in mental rotation. In the second experiment, the maze had proximal cues, in the form of landmarks on the circumference of the virtual pool, and gender-based differences in navigational performance significantly diminished. Under these changed environmental conditions, mental rotation ability did not correlate with performance in the VMWM, suggesting that given proximal cues, the need for mental rotation diminishes. Differences between videogame novices and experts also decreased when proximal cues were provided. Females in particular obtained more discernible benefits from videogame experience. Together, these experiments reveal how the spatial abilities and strategies used to solve the Morris maze task vary with environmental design. Given the structural similarities between the virtual maze and videogame environments, these results offer insight into how spatial experience gained through videogame playing can affect aspects of spatial cognition, and can help identify design elements that contribute to their improvement.

1. Why (and how) spatial abilities matter

> *"[Spatial ability] may be the largest known untapped source of human potential. And yet, no admissions directors I know of are looking at this, and it's generally overlooked in school-based assessments."*
>
> David Lubinski as quoted in Clynes (2016)

Advancing our understanding of spatial abilities matters: it can contribute directly to the advancement of educational and socio-economic equity. In her review of gender, spatial abilities and wayfinding research, Carol Lawton (2010) stresses that "Spatial abilities …show the clearest evidence of gender differences in cognition" (p.317), and she cites over 50 years of research that might suggest gender difference in spatial abilities as a reason for women's under-representation in scientific and technical fields (Lawton, 2010). Although these findings could be (and have been) used to justify female under-representation, Lawton also indicates that "…research on gender and spatial cognition can be used to inform methods of improving spatial abilities" (p. 317) which is, she says, "perhaps its greatest value" and she recommends that "efforts should be made to incorporate spatial training into early education before gender-based stereotypes about spatial abilities have been learned." (Lawton, 2010, p. 332). To contribute to realizing that goal is the motivation for and the purpose of the research reported here.

Spatial abilities have been studied using a variety of psychometric tests, though without consensus on how they should be defined and classified (Voyer, Voyer, & Bryden, 1995). Generally, spatial ability has been defined as "the ability to generate, retain, retrieve, and transform well-structured visual images" (Lohman, 1994); the ability to manipulate visual images mentally (Burnett, Lane, & Dratt, 1979); and the ability to create "a mental representation of a two or three-dimensional structure and then assessing its properties or performing a transformation of the representation" (Carpenter & Just, 1986, p. 221). As Voyer et al. (1995) point out, these definitions are vague and overlook the diverse processes underlying each spatial task, making it difficult to replicate and compare studies.

Moreover, tests that purport to measure one component (such as spatial visualization) also involve elements of other processes (such as mental rotation), making it difficult to classify results under a single category (Voyer et al., 1995). When discussing the issue of spatial ability, they argue, caution is required, and the demands of each spatial task should be carefully examined, because spatial ability is not a unitary concept but a general category for a set of diverse and often

☆ This research was supported by the Social Sciences and Humanities Research Council of Canada.

* Corresponding author.

E-mail addresses: suzanne.decastell@uoit.ca (S. de Castell), hlarios@sfu.ca (H. Larios), jjenson@edu.yorku.ca (J. Jenson).

https://doi.org/10.1016/j.actpsy.2019.102895

Received 5 February 2019; Received in revised form 3 June 2019; Accepted 16 July 2019

Available online 01 August 2019

interrelated abilities.

Mental rotation is one spatial ability that has received widespread attention in the literature, with > 5000 studies conducted (Ganis & Kievit, 2015) across a wide variety of contexts, including wayfinding (Malinowski, 2001), sign language (Talbot & Haude, 1993), object recognition (Tarr, 1995), and virtual environments (Parsons et al., 2004). Shown to be significantly associated with the ability to solve types of problems encountered in STEM fields, and with engagement in STEM careers (Humphreys, Lubinski, & Yao, 1993; Lubinski, Benbow, & Kell, 2014; Uttal & Cohen, 2012; Wai, Lubinski, & Benbow, 2009), mental rotation has been studied extensively in relation to sex differences, showing a consistent male advantage (Voyer et al., 1995).

Uttal and Cohen (2012) indicate that research on the relationship between STEM education and spatial abilities has focused on spatial visualization, seen as “the process of apprehending, encoding and mentally manipulating three-dimensional spatial forms” (Carroll, 1993 as cited in Uttal & Cohen, 2012). From this perspective, they consider mental rotation a form of spatial visualization. In Shepard and Metzler's (1971) original test, participants are shown pairs of abstract figures made up of connected cubes. The objects appear rotated in relation to each other, and the task is to determine whether the pairs are the same or different. The fact that error rates and response times increase linearly with angular disparity is taken as evidence that participants are mentally visualizing and rotating the objects to solve the task (Shepard & Metzler, 1971).

Because of the difficulty of disambiguating mental rotation and spatial visualization, the terms are often used interchangeably, and although some researchers have classified these as separate spatial abilities (Linn & Petersen, 1985), it is difficult to imagine mental rotation of an image without spatial visualization.

While the precise abilities that lead to success in STEM disciplines have not been thoroughly delineated, mental rotation appears to be a key factor, explicitly linked to success in a variety of STEM disciplines: biology (Russell-Gebbett, 1985), medicine (Hegarty, Keehner, Cohen, Montello, & Lippa, 2007), geology (Orion, Ben-Chaim, & Kali, 1997), chemistry (Stieff, 2007), physics (Kozhevnikov, Motes, & Hegarty, 2007) and engineering (Lajoie, 2003). For these reasons we have relied on tests of mental rotation as paradigmatic in assessing the kinds of spatial abilities conducive to success in STEM education and STEM careers.

2. Videogames, navigation, and spatial cognition

Three-dimensional (3D) virtual environments (VEs) are an increasingly useful tool for the visualization, manipulation, and processing of information across a variety of domains such as education, science, and engineering. In VEs, information can be displayed in a graphically realistic and intuitive manner, allowing users to navigate the environment using skills associated with spatial cognition (Bowman, Kruijff, LaViola, & Poupyrev, 2004). To better understand the basis of this strong association of spatial abilities with success in science, technology, engineering, and mathematics (STEM) (Humphreys et al., 1993; Lubinski et al., 2014), VEs can be harnessed to investigate the specific factors underlying spatial reasoning, to predict aptitude for STEM education and occupations, and to design interventions that improve spatial skills and abilities.

Spatial skills training has been shown to be effective, long-lasting and transferable to contexts outside the training regime (Sorby, 2009). Unfortunately, spatial skills training does not yet enjoy widespread acceptance and inclusion in educational curricula, and there remains much yet to learn about what and how training can lead to improvements in spatial skills. One promising line of research shows that playing videogames can improve spatial cognition and perception (Boot, Blakely, & Simons, 2011; De Lisi & Wolford, 2002; Feng, Spence, & Pratt, 2007; Green & Bavelier, 2003, 2007, 2012 and Okagaki & Frensch, 1994). Researchers have found that playing abstract games like Tetris (Boot, Kramer, Simons, Fabiani, & Gratton, 2008; Okagaki & Frensch, 1994) and first-person action videogames (Feng et al., 2007; Quaiser-Pohl, Geiser, & Lehmann, 2006) is associated with improvements in mental rotation. Questions remain about whether improvement in a specific context (such as improving mental rotation through videogames) is transferable to other contexts. Uttal and Cohen's (2012) meta-analysis found evidence of transfer in some studies (e.g., Sorby, 2009Sorby & Baartmans, 1996, 2000) but not in others (Sims & Mayer, 2002).

Videogames have become increasingly popular vehicles for education and training, with the strongest evidence of their effectiveness being to improve spatial abilities (Feng et al., 2007; Greenfield, 2009; Oei & Patterson, 2013). To the best of our knowledge, little is yet known about what kinds of games and game elements might be responsible for these improvements. Videogame environments like those found in 3D action games are relatively complex, and they require players to orient themselves in a variety of locales such as buildings, city streets, and natural settings. It is conceivable that navigation in these graphically and procedurally realistic environments elicits the same spatial abilities (e.g., mental rotation) required in real-world navigation. Results from several studies in environmental cognition have shown that small-scale spatial abilities (e.g., mental rotation and perspective-taking) are associated with navigation in large-scale settings (Bryant, 1982; Hegarty, Richardson, Montello, Lovelace, & Subbiah, 2002; Malinowski, 2001; Waller, 2000). Like real-world environments, videogame settings pose several experimental challenges for researchers aiming to investigate how videogames improve spatial abilities. The gaming environment is richly detailed with many graphical elements that can potentially be used as navigation aids- landmarks, geometry, textures, colours, and variety of dynamic elements. Moreover, players engage in multiple tasks such as finding hidden objects, shooting enemies, searching for locations, swimming, running, and so forth. These complexities make it difficult to isolate and control the precise factors that contribute to enhancing spatial abilities. To investigate one such set of factors in a more controlled laboratory setting, we developed a ludic virtual environment based on the Morris Water Maze paradigm, which is described in the following section.

3. The (virtual) Morris Water Maze

Past research indicates that males and females differ in wayfinding strategies and associated spatial orientation skills. Specifically, males and females have been shown to rely on different types of sensory cues when navigating and orienting themselves. In their review of the literature, Coluccia and Louse (2004) point out that sex differences in landmark use have been found in tasks related to map sketching (McGuinness & Sparks, 1983), verbal description of routes (Brown, Lahar, & Mosley, 1998; Dabbs, Chang, Strong, & Milun, 1998; Miller & Santoni, 1986; Schmitz, 1997; Ward, Newcombe, & Overton, 1986), and self-report questionnaires (Lawton, 1994; O'Laughlin & Brubaker, 1998; Pazzaglia & Cornoldi, 1999). These studies have also found that males tend to rely on configurational and global (distal) properties of the environment (e.g., cardinal points, metric distances, and geometric cues). The exact cognitive mechanisms behind these sex differences in spatial orientation strategies are currently unknown.

Research with rodents has found male rats tend to rely on configurational cues when completing orientation tasks in a maze (Foreman, 1985; Margueles & Gallistel, 1988; Suzuki, Augerinos, & Black, 1980; Williams, Barnett, & Meck, 1990; Williams & Meck, 1991). Extending this research paradigm to humans, sex differences have also been found using virtual analogues of the Morris Water Maze (Astur, Ortiz, & Sutherland, 1998; Mueller, Jackson, & Skelton, 2008; Newhouse, Newhouse, & Astur, 2007; Sandstrom, Kaufman, & Huettel, 1998). Originally, the Morris Water Maze (MWM) was developed to study spatial learning and memory in rodents (Morris, 1984). More than 2000 neuroscience experiments have since been conducted in which the

MWM methodology has been implemented, demonstrating the scientific utility of this tool (D'Hooge & De Deyn, 2001). The MWM consists of a circular pool within a rectangular room with different types of visual cues. A typical test consists of placing a platform in the water and allowing the subject (usually a rat) to swim freely in search of the platform. Researchers measure variables such as how long it takes the subject to find the platform (search latency) and the amount of time searching the correct area (dwell time). These and other measures are taken as an indication of the strength and accuracy of spatial learning and working memory (D'Hooge & De Deyn, 2001).

Research on humans using the MWM is limited because of ethical concerns to health and safety, cost, and procedural complexities. To overcome the limitations posed by the physical MWM, virtual versions of the MWM (VMWM) have been developed to investigate human spatial cognition. Sandstrom et al. (1998), for example, report that "When tested in an environment analogous to the MWM often used in the study of rodent spatial navigation, humans performed in a fashion similar to rats by showing reduced latencies to the target with repeated trials and a dependence on distal environmental cues" (p. 357).

Sandstrom et al. (1998) found that female participants' performance in the VMWM task was better when distal landmarks were available and could be used to accurately predict the location of the hidden platform. However, their performance was adversely affected when distal landmarks were unreliable or not available. Males, on the other hand, did not show major deficiencies in performance in similar circumstances because they seem to be able to switch strategies by relying on distal landmarks or geometrical-configurational cues when required. Sandstrom et al. (1998) did not investigate whether the manipulation of proximal landmarks had a similar effect on males' and females' search strategies. Livingstone and Skelton (2007) investigated the use of distal and proximal cues by traumatic brain injury (TBI) survivors that completed the VMWM navigation task. They found that TBI survivors were able to navigate to a target using proximal landmarks but showed deficiencies in performance when the landmarks were absence. The researchers did not investigate sex differences in performance. Nevertheless, their study shows the utility of the VMWM in examining issues relevant to the use of proximal and distal cues during spatial learning and memory.

While past research has focused on the strategies used to complete the VMWM task, relatively little research exists on the spatial abilities employed during VMWM navigation. Nevertheless, a significant correlation between performance in the VMWM and mental rotation has been observed in a number of studies (see Astur, Tropp, Sava, Constable, & Markus, 2004; Driscoll, Hamilton, Yeo, Brooks, & Sutherland, 2005; Sneider et al., 2015). Given that improvements in mental rotation have also been observed after training in 3D action videogames (Feng et al., 2007), it is conceivable given the similarities between gaming environments and the VMWM, that we can investigate the precise elements that lead to improvements in mental rotation abilities.

Scouring through the literature on uses of the VMWM revealed an array of differences in the ways these experimental environments have been designed. It was surprising to find little explicit description of the specifications of the VMWM used in each study, and little attention to the value of developing a 'standard' model and practices. After much deliberation, we built our own maze (described below) from the specifications provided by Mueller et al. (2008). We sought to determine the utility of the VMWM in investigating how males and females with varying degrees of videogame expertise orient themselves in the presence and absence of proximal cues. As well, we investigated the role of mental rotation in spatial learning and memory in completing the maze task.

4. Research aims

We examined spatial cognition in a VMWM to study spatial learning and memory, and to see how well our findings align with gender differences found in prior research. We also sought to identify the impact of spatial experience (specifically videogames) on task performance. We expected to find sex differences in VMWM performance favouring males and videogame experts. Previous VMWM research has found a gender difference in uptake and use of distal cues (Mueller et al., 2008; Sandstrom et al., 1998), but it is not clear how proximal landmarks affect navigation performance in this measure of spatial learning and memory. To see how these elements in a virtual environment affected performance, we developed two versions of the VMWM, one containing proximal information consisting of five landmark-type cues around the pool wall, and the other lacking proximal landmarks. We conducted experiments with each version of the maze. We expected that when proximal cues were not visible, participants would rely on distal information (e.g., room walls and geometric information), leading to a large and significant difference in performance between males and females, as well as video game experts and novices. We also predicted an improvement in performance with the addition of proximal cues, minimizing sex and expertise differences. We administered Peters, Laeng, Latham, and Jackson's (1995) version of the Vandenberg and Kuse (1978) mental rotation test, and hypothesized that male and videogame experts would exhibit significantly higher mental rotation scores than females and videogame novices. Lastly, examining the correlation between mental rotation and search latency in the VMWM, we predicted that higher mental rotation abilities would be associated with better search performance. With some key exceptions (particularly those related to videogame experience), these predictions held, and we discuss their importance for the study and training of spatial skills.

5. Methods

5.1. Participants

Eighty-two participants from Simon Fraser University volunteered to participate at the School of Interactive Arts and Technology in Surrey, BC, Canada. Participants were recruited using a variety of methods including posters, email, and an online participant recruiting system accessible to university students. Forty-five of the participants (16 males and 29 females) completed experiment 1, using the version of the maze that contained no visible proximal cues. Of these 45 participants, 25 were assessed as videogame novices (5 male and 20 female), and 20 as experts (11 male and 9 female). Thirty-seven participants (16 males and 21 females) completed experiment 2, using the version of the maze that contained proximal cues. They were further grouped by videogame expertise into 18 novices (4 male and 14 female), and 19 experts (12 male and 7 female). The participants received a $25 honorarium. Each experimental session lasted approximately 1 h.

The mean age of the sample population in the first experiment was 21.8 years (SD = 2.4) with a range of 18–28 years. In the second experiment, the mean age was 21.2 year (SD = 2.5) with a range between 18 and 28 years.

Participants completed a demographics questionnaire (Appendix A) and videogame survey (Appendix B), which were administered towards the end of each experimental session to avoid priming responses, behaviour, and performance (Guizzo, Moe, Cadinu, & Bertolli, 2019). Further, Boot et al. (2011) caution, "One possible factor that could lead to the spurious conclusion of gaming benefits on cognition is differential expectations for experts and novices." In the demographics questionnaire, participants indicated whether English was their first or second language; they considered themselves left or right handed; had taken courses in videogames, 3D graphics, and animation; and if they were involved in sports. In the Videogame survey, participants answered questions regarding their videogame habits, which were used to assess their level of expertise.

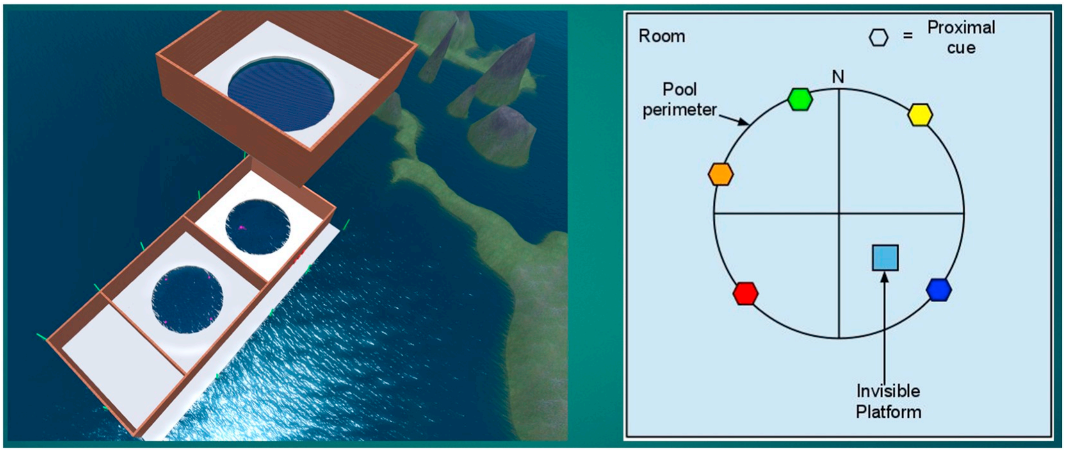

Fig. 1. Overhead view of the VMWM and a schematic representation.

5.2. Stimuli

The virtual maze consisted of a circular pool with a 40 virtual-meter diameter, housed within a rectangular room measuring 60 × 60 × 17.5 m (Fig. 1). The pool was filled with water to a depth of 20 m and was bounded by a 1.5 m high wall. The practice pool had a similar design to the main testing pool, but did not contain any water, because the intent was to familiarize the user with the navigation controls and general movement within the environment.

In the "Distal Cues" experiment, no proximal landmark cues were added to the pool environment. In the Proximal Cues experiment, we distributed five proximal cues (images of marine creatures) around the perimeter of the pool wall (Fig. 2).

The platform was a bright pink circle (Fig. 3) with a 1.5 m diameter, located at a distance of 7 m from the pool wall. Participants navigated the maze using a rat avatar as shown in Fig. 2.

5.3. Procedure

5.3.1. Behavioural testing

At the start of each session, the experimenter read the study overview and provided participants with a consent form. The experimenter answered questions and debriefed participants at the end of the session. Participants were told that they were going to play a game that required them to find the location of a platform located in the pool. They were instructed to use the arrow keys to navigate in the pool, but they were asked not to use the 'back' arrow, as in traditional MWM experiments animal subjects do not swim backwards. Testing proceeded with a series of trials adapted from Mueller et al. (2008): Practice, visible platform (VP), hidden platform (HP), and probe trials (Table 1).

5.3.2. Tests and questionnaires

After the VMWM task, participants completed Peters et al. (1995) redrawn Vandenberg and Kuse Test of Mental Rotation. In this 24-item test, participants were presented with a target figure and four stimulus figures. Two of the stimulus figures were rotated versions of the target figure. Participants needed to find both figures and mark them with an X to receive a point. Once participants completed the mental rotation test, they filled out a videogame experience survey, kindly provided to us by Dr. Walter R. Boot from Florida State University (personal correspondence).

5.3.3. Measures

We recorded the start and end times (in seconds) of each VP and HP

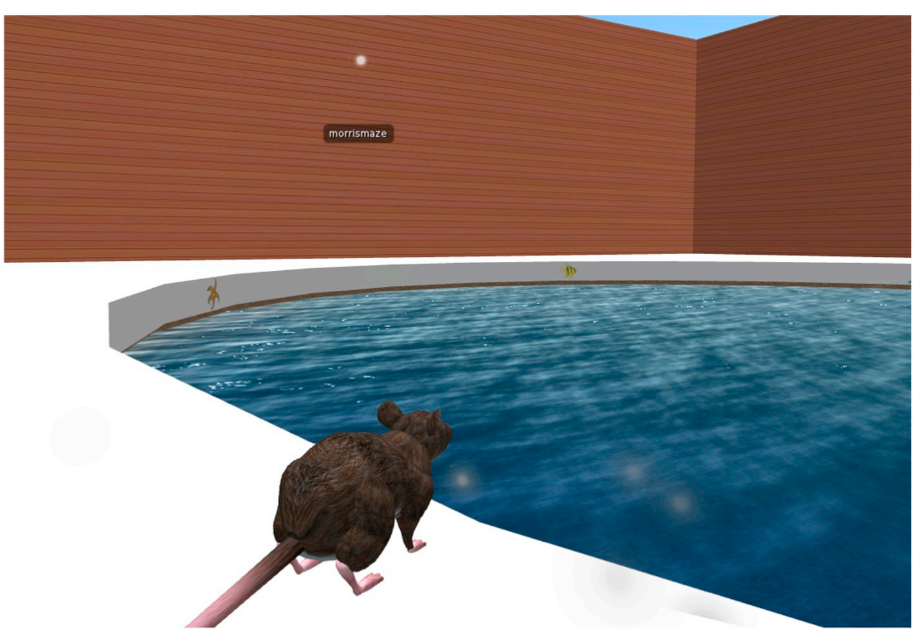

Fig. 2. The VMWM with images placed on the pool wall.

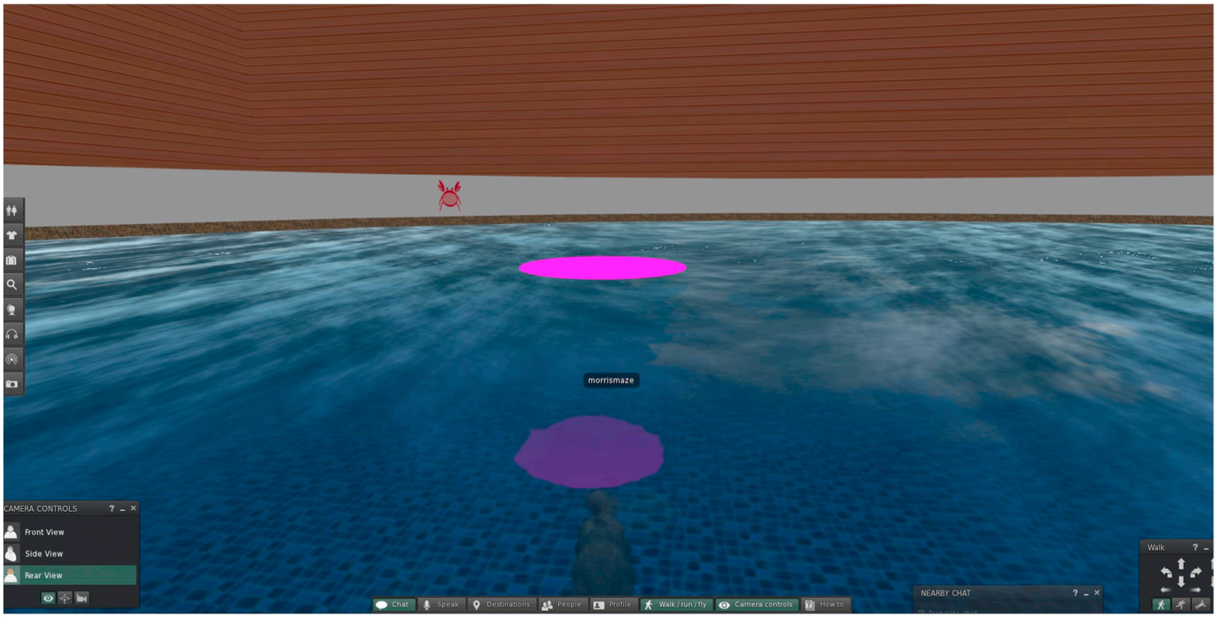

Fig. 3. Platform area in the VMWM. (For interpretation of the references to color in this figure, the reader is referred to the online version of this chapter.)

trial. The duration between start and end times is the search latency (SL). On the last trial, called the probe trial, we analysed the amount of time that participants spent searching for the platform on the correct quadrant where the platform had been located in the preceding 10 HP trials. This measure is known as quadrant dwell time.

Participants were grouped into videogame novices and experts based on their responses to several survey questions about their gaming habits with 3D videogames. Each question contained a 5-point scale on the frequency of gameplay: Never, Seldom, Sometimes, Frequently, and Often. A participant was considered a videogame novice on a video game or category if he/she answered ‘never’ or ‘seldom’ to any of the relevant survey question. The participant was considered an expert if he/she answered ‘frequently’ or ‘often’. If a participant answered ‘sometimes’, then we examined the rest of their survey answers to determine the quality (e.g., type of games they preferred) and quantity of play (e.g., number of hours per week). As per Boot et al. (2008) “Participants were considered experts if they played seven or more hours of video games per week for the past two years. Non-gamers were selected such that they played video games one hour a week or less.”

The key survey items used to assess videogame experience are listed below.

How frequently do you do the following:

1. Play DOOM, Call of Duty, Halo, Battlefield or similar first-person shooters?
2. Play Medal of Honor (any version, e.g. Medal of Honor Allied Assault, Medal of Honor Pacific Assault, etc.…)?
3. Play Final Fantasy or similar role-playing video games?
4. Play simulator video games (e.g., flight simulator or racing simulator games)?
5. Play Grand Theft Auto or similar action/platform games?
6. Play sports video games (e.g., NBA Live, Madden NFL, FIFA Soccer, SSX, Tony Hawk, or similar games)?

5.4. Data analysis

Search Latencies (SLs) were analysed using a two-factor mixed-design ANOVA (sex × VG expertise) with *trial* as a repeated measure. A *Huynh-Feldt* correction was applied if the data violated the sphericity assumption. The data from the first HP trial were excluded from the analysis because participants had not yet seen the location of the platform in this trial, making this an exploratory ‘scouting’ of the environment which could provide no salient information on their search strategies. To improve the efficiency of the analysis by attaining an adequate signal-to-noise ratio, the last nine HP trials were collapsed into three blocks of three trials each. Separate ANOVAs were used for the VP (×4) and HP blocks (×3) of trials. The quadrant dwell times in the probe trial were analysed with a sex × expertise ANOVA.

Scores of mental rotation were examined with independent samples *t*-tests. Correlations between spatial abilities, search latency, and quadrant dwell time were examined with Kendall's τ non-parametric correlations.

For all statistical tests, a significance level of 0.05 was used.

Table 1
Description of search trials in the VMWM (Adapted from Mueller et al., 2008).

Type	#	Trial description
Practice	1	The participant navigated an empty pool to familiarize himself/herself with the keyboard controls and the environment.
Visible Platform (VP)	4	The participant teleported into one of the four randomly determined starting locations (NSEW) and searched for the platform, which was clearly visible in one of the four quadrants. On subsequent trials, the participant starting location was randomized and the platform changed location around all four quadrants. When the participant found the platform, the avatar was held in place and a male voice said, “Congratulations! You have found the platform and escaped from the waters.”
Hidden Platform (HP)	10	The participant began testing in a pseudo-randomly determined starting location and had 3 min to find the hidden platform, which was always located in the SE quadrant of the pool. If the participant found the platform they received the same message as in the VP trials. If they failed to find it, the platform appeared and a voice said, “Time has expired. Please swim to the platform.”
Probe	1	On the last trial, the platform was in the same location as in the previous trial, but it was inactive for 20 s. After 20 s, the platform became visible and the participant was instructed to move to the platform.

6. The distal cues experiment – results and discussion

In this experiment, the VMWM did not have any proximal landmarks that participants could rely on to find the platform. This design was intended to increase their reliance on distal cues. We predicted that males would outperform females by obtaining shorter latencies in the hidden platform trials and higher dwell times in the probe trial. Similarly, we also predicted that videogame experts would demonstrate better performance than novices. In mental rotation, we hypothesized that males would outperform females and experts would outperform novices. Lastly, we predicted that mental rotation scores would correlate with maze performance. Although we did not make specific predictions on the role of sex and videogame experience on mental rotation, we further examined the correlations based on the participants' sex and expertise.

6.1. Results

6.1.1. Search latencies

6.1.1.1. Visible platform trials. The participants' estimated marginal (EM) mean SLs with their corresponding Standard Error (SE) are reported in Table 2.

Although overall participant performance appeared to decrease over the course of trials, the main effect of trial was not significant [$F(2.79, 114.22) = 2.54$, $p = .06$]. There was no significant interaction of trial with sex [$F(2.79, 114.22) < 1$, $p > .05$]. Fig. 4 shows that males consistently obtained lower SLs than females on the last three trials, but the main effect of sex was not significant [$F(1, 41) = 3.10$, $p = .09$].

As illustrated in Fig. 5, experts consistently obtained shorter latencies across the four VPTs. The interaction effect of trial and expertise, however, was not significant [$F(2.79, 114.22) < 1$, $p > .05$]. The main effect of VG expertise was not significant either [$F(1, 41) = 2.65$, $p > .05$].

The EM mean latencies grouped by sex × expertise can be found in Table 6 and Table 7 in Appendix C. Fig. 6 illustrates VPT latencies by male novices and experts separate from those of female novices and experts (Fig. 7). The interaction effect of trial, sex, and expertise was not statistically significant [$F(2.79, 114.22) < 1$, $p > .05$] and neither was the interaction effect of sex and expertise [$F(1, 41) < 1$, $p > .05$].

6.1.1.2. Hidden platform trials. Analysis of the HPT blocks yielded a significant main effect of trial block [$F(1.94, 79.72) = 4.44$, $p < .05$], which indicated that participants obtained lower SLs as the trials proceeded. Table 3 and Fig. 8 show that males and females obtained relatively similar SLs in the first block of trials, but as the trials continued, males found the hidden platform faster than females. The interaction effect of trial block and sex was significant [$F(1.94, 79.72) = 6.49$, $p < .01$]. Post hoc LSD pairwise comparisons showed that males significantly differed from females in the second ($p < .05$) and third ($p < .01$) trial blocks. The main effect of sex was also significant [$F(1, 41) = 5.44$, $p < .05$] with males showing overall lower adjusted SLs.

The SLs for VG experts and novices are illustrated in Fig. 9. Although the adjusted means show that experts obtained faster SLs across the first blocks of trials, the interaction effect of trial and expertise was

Table 2
Adjusted mean search latencies (seconds) and standard errors (SE) in the visible platform trials.

Visible platform trial	Females (SE)	Males (SE)	Experts (SE)	Novices (SE)
1	11.3 (0.84)	10.8 (1.13)	9.8 (0.94)	12.3 (1.05)
2	10.6 (0.81)	9.2 (1.08)	9.6 (0.90)	10.2 (1.01)
3	10.4 (0.70)	9.0 (0.93)	9.1 (0.78)	10.4 (0.87)
4	9.8 (0.52)	7.9 (0.69)	8.7 (0.576)	9.1 (0.64)
Overall	10.5 (0.44)	9.2 (0.59)	9.28 (0.49)	10.5 (0.55)

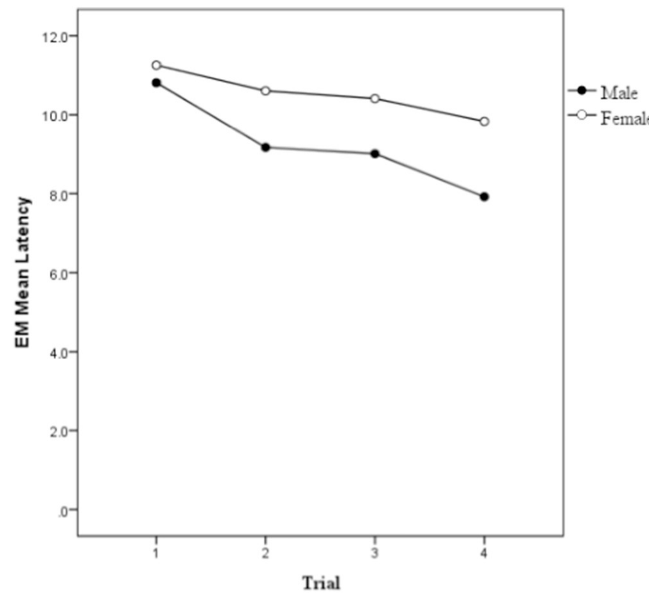

Fig. 4. Males and Females adjusted mean search latencies (s) in the visible platform trials.

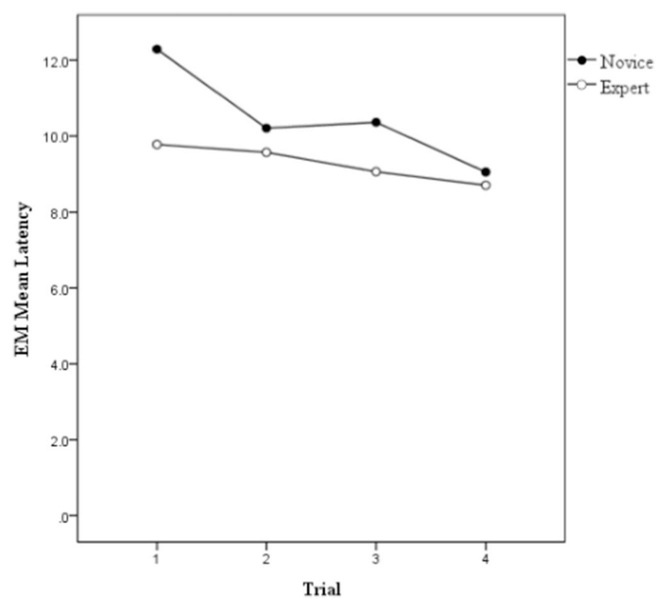

Fig. 5. VG experts' and novices' adjusted mean search latencies (s) in the visible platform trials.

not significant [$F(1.94, 79.72) = 1.42$, $p > .05$] and neither was the main effect of expertise [$F(1, 41) = 1.36$, $p > .05$].

Adjusted mean SLs grouped by participants' sex × expertise are reported on Table 8 and Table 9 in Appendix C and shown in Fig. 10 and Fig. 11. The analysis did not yield a significant effect of trial block with sex × expertise [$F(1.94, 79.72) = 1.14$, $p > .05$]. Further, the interaction effect of sex × expertise was not significant [$F(1, 41) < 1$, $p > .05$].

6.1.1.3. Probe trial. In the probe trial, a main effect of sex showed that males ($M = 12.0$, $SE = 1.61$) spent significantly longer time on the platform quadrant than females ($M = 6.4$, $SE = 1.20$) [$F(1, 41) = 7.9$, $p = .01$]. There was no significant main effect of expertise [$F(1, 41) < 1$, $p > .05$] or interaction effect of sex × expertise [$F(1, 41) < 1$, $p > .05$].

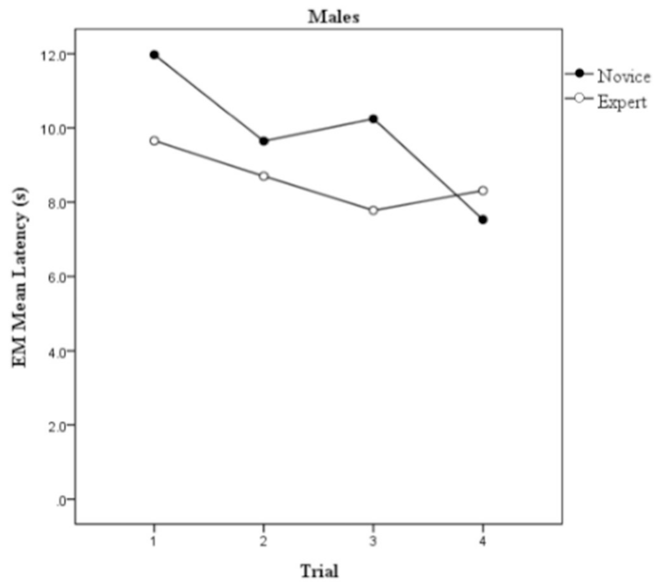

Fig. 6. Male VG experts' and novices' adjusted mean search latencies (s) in the visible platform trials.

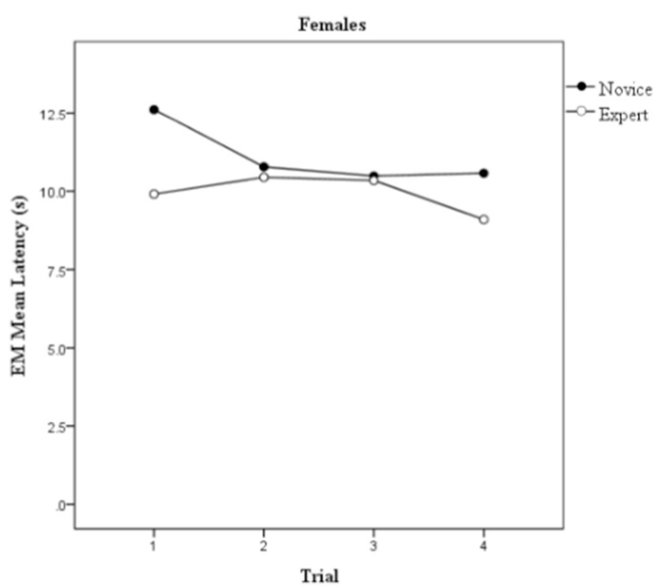

Fig. 7. Female VG experts' and novices mean search latencies (s) in the visible platform trials.

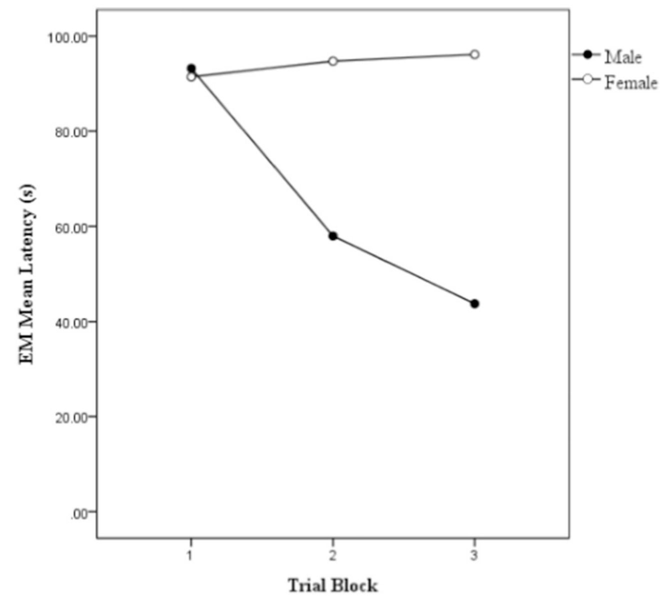

Fig. 8. Males' and females' mean search latencies (s) in the HPT blocks.

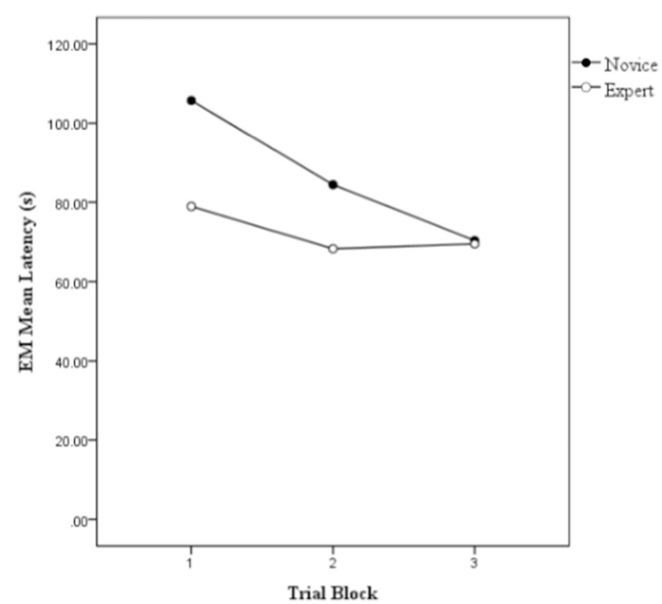

Fig. 9. VG experts' and novices' adjusted mean search latencies (s) in the HPT blocks.

Table 3
Adjusted mean search latencies (s) and standard errors (SE) in the HPT blocks.

HPT block	Females (SE)	Males (SE)	Experts (SE)	Novices (SE)
1	91.5 (8.83)	93.2 (11.87)	79.0 (9.89)	105.7 (11.00)
2	94.8 (8.97)	58.0 (12.06)	68.3 (10.05)	84.4 (11.18)
3	96.15 (9.68)	43.7 (13.00)	69.5 (10.84)	70.4 (12.05)
Overall	94.12 (7.46)	65.0 (10.02)	72.3 (8.35)	86.8 (9.29)

6.1.2. Mental rotation test - differences and correlations

6.1.2.1. Sex. Analysis of the Mental Rotation test scores (Fig. 12) yielded a statistically significant difference between males (14.4 ± 0.95) and females (9.4 ± 1.01) $t(43) = 3.31$, $p = .001$, $r = 0.45$.

There was a statistically significant *negative* correlation between MR scores and mean search latency in the HP trials, $\tau = -0.26$, $p < .01$. The higher the MR score, the faster the participants reached the platform. Separately, however, neither males' MR scores, $\tau = -0.21$, $p > .05$, nor females' scores, $\tau = -0.10$, $p > .05$, correlated significantly with SL in the HP trial blocks. On the probe trials, MR scores *positively* correlated with search dwell time on the platform quadrant, $\tau = 0.24$, $p < .01$. That is, better MR scores were associated with more time spent on the correct quadrant, which is an indication that the participant has learned the general location of the platform. Analysed by sex, females' MR scores were not significantly correlated with dwell time on the platform quadrant, $\tau = 0.10$, $p > .05$. Similarly, males' MR scores did not significantly correlate with quadrant dwell time, $\tau = 0.18$, $p > .05$.

The results showed that SL in the VPTs negatively correlated with MR scores $\tau = -0.21$, $p < .05$. In other words, MR predicted performance in the VPTs, suggesting a spatial component to the VPTs.

6.1.2.2. Videogame expertise. On the MR test (Fig. 12), experts (13.9 ± 1.23) scored significantly higher than novices (9.0 ± 0.88) *t*

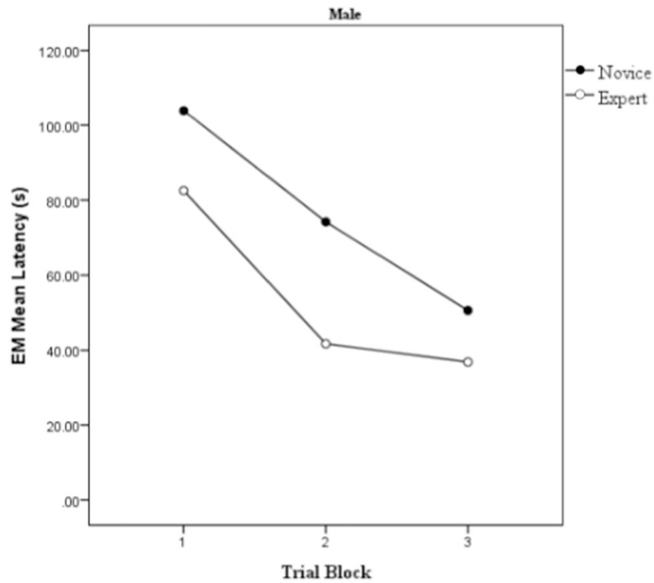

Fig. 10. Male VG experts' and novices' adjusted mean search latencies (s) in the hidden platform trials.

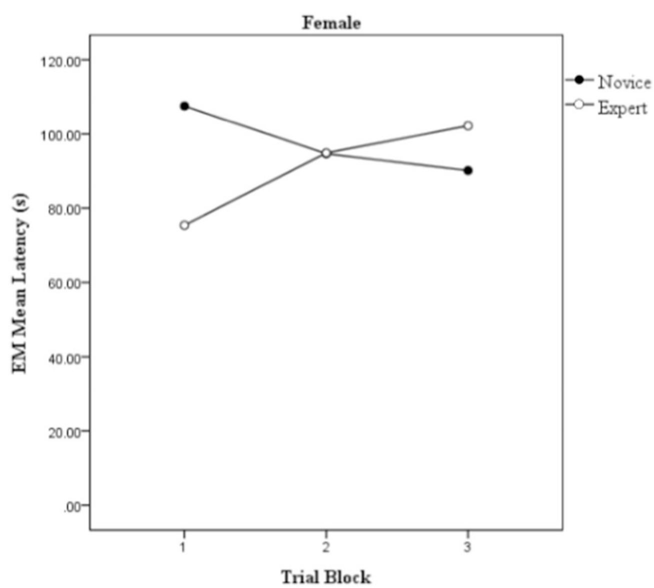

Fig. 11. Female VG experts' and novices' adjusted mean search latencies (s) in the hidden platform trials.

(43) = 3.33, $p = .001$, $r = 0.45$. Novices' MR scores were significantly correlated with SL in the HP trials, $\tau = -0.26$, $p < .05$. The MR scores for experts, on the other hand, were not correlated with SL, $\tau = -0.13$, $p > .05$. On the probe trial, a similar pattern emerged, as the novices' MR scores significantly correlated with search dwell time on the platform quadrant, $\tau = 0.34$, $p = .01$. The experts' MR scores did not significantly correlate with quadrant dwell time, $\tau = -0.43$, $p > .05$. Although we expected the opposite result, this finding aligns with Uttal and Cohen's (2012) observation that novices tend to rely on spatial abilities in the early stages of learning, but these abilities become less important as domain expertise increases.

Although, we did not conduct a fine-grained analysis of sex × expertise in mental rotation, the descriptive results show the tentative finding that videogame experts outperformed novices across both sexes (Fig. 13).

6.2. Discussion

In this first experiment, participants conducted a search task in a version of the virtual environment that did not feature any proximal landmarks. Consistent with past research, a significant sex difference was found, showing that males were faster than females in reaching the hidden platform. As well, males spent a significantly higher proportion of time searching in the correct platform quadrant during a probe trial, which is considered by some researchers as "the single best measure of spatial strategy" (Newhouse et al. 2007). A test of mental rotation also replicated previous findings demonstrating a male advantage. Moreover, high scores on the MRT were correlated with better performance on the maze, though not when males and females scores were analysed separately.

Acquisition curves for the HPTs showed that in the early block of trials performance between males and females was comparable. However, as the trials proceeded, males search times decreased while those of females did not show this pattern. This result suggests a stronger acquisition and cognitive consolidation of the platform location across repeated trials for males but not females. In our review of the *virtual* maze literature, specifically, we found only one other report of this outcome. Newhouse et al. (2007) tested pre-pubertal children in a VMWM and found that boys but not girls significantly decreased their escape latencies across blocks of trials. Their results suggested that "girls did not substantially improve their latency over training blocks" (Newhouse et al., 2007). While they did not propose an explanation for this finding, the researchers indicated that overall sex differences in VMWM performance could be attributed to the use of "Euclidean versus Landmark" strategies and abilities. A Euclidean strategy is defined as one in which a participant attends to and utilizes non-distinct (geometric) cues that provide distance and direction information (Dabbs et al., 1998). A landmark strategy is one in which salient and distinct topographic features are used as navigation aids (see Dabbs et al., 1998; Ward et al., 1986). Newhouse et al.'s (2007) hypothesis that sex differences in VMWM performance might be the result of males' ability to use both Euclidean and landmark strategies during navigation and females' preferential reliance on landmark strategies alone has found some support by other researchers (e.g., Moffat, Hampson, & Hatzipantelis, 1998; Sandstrom et al., 1998).

Our version of the VMWM for the first experiment had no distinct landmark cues. After completing the task, participants were asked what strategy they used to find the platform. Among the responses, the participants mentioned that they:

- Estimated the location of the platform by looking the height of the pool wall when they stood on different areas of the pool
- Tried to align themselves with a small visible portion of the sky above the room walls
- Tried counting small tiles at the bottom of the pool while standing on specific locations
- Aligned the avatar with the corners of the room while moving back and forth
- Circled the pool at a specific distance from the wall
- Attended to the flow of the water and light reflections on it
- Looked at the differences in luminance on the walls
- Used the corners of the monitor

Often, participants abandoned these strategies and could not say with certainty how they learned the location of the platform. This shows a limitation with verbal reports and how caution is necessary when interpreting them, since participants may be unconscious of and unable to describe the sensory, perceptual, and cognitive processes underlying their own performance. Participants may also not be able to accurately recall strategies, particularly if they were not consciously aware of what strategies they used. Nevertheless, the participants' responses give an insight into possible strategies that could be used when

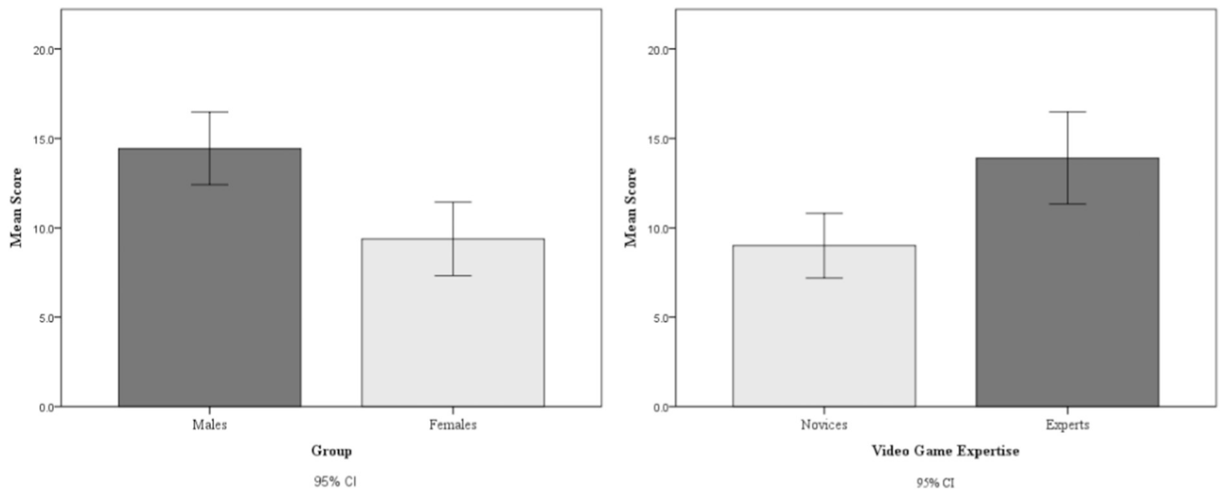

Fig. 12. Mental Rotation test scores grouped by sex and expertise separately.

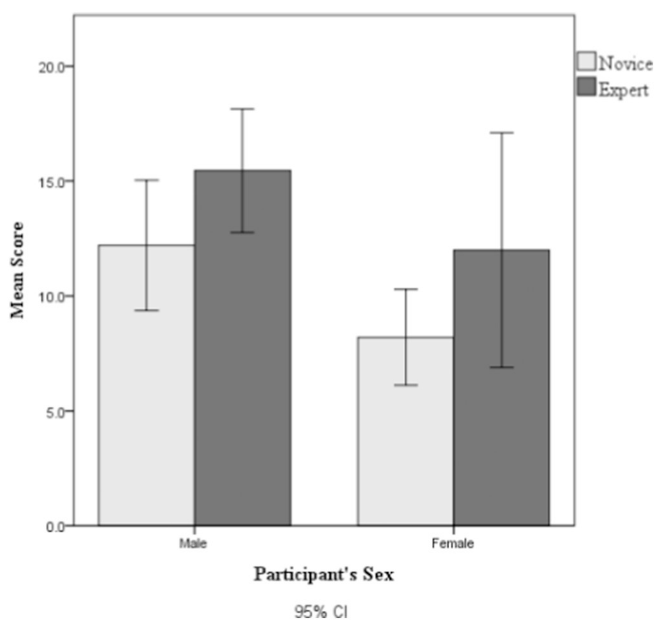

Fig. 13. Mental rotation scores grouped by sex × expertise.

completing the VMWM task in an environment that lacks salient proximal cues. Although participants' responses do not permit a quantitative analysis that would give us insight into which strategy was the most used and whether gender differences in strategy were evident, they do illuminate the strategies that were most salient in the participants' minds.

Although the VMWM is used to measure spatial learning and memory, Newhouse et al. (2007) point out that participants need not solely rely on spatial learning and memory to learn the location of the platform in the HPTs. For instance, participants could–and in our study reported that they did– use non-spatial strategies (e.g., circling pool) to locate the hidden platform. Given the probability that SLs do not accurately reflect the use of a spatial strategy, the maze task contains a probe trial during which the platform is inactive. A participant who has learned to spatially locate the platform will demonstrate a bias for the quadrant in which the platform is located and spend 50% or more search time in this platform quadrant (D'Hooge & De Deyn, 2001). In the present experiment, male participants spent 62% of their search time in the platform quadrant. Females, on the other hand, spent approximately 30% of their search time there, which is close to chance. This result is evidence of a sex difference in spatial learning and memory, and it is consistent with past research (e.g., Astur et al., 2004; Mueller et al., 2008; Newhouse et al., 2007). Unlike past research, we cannot attribute this result to the differential use of landmark cues because no such cues were visible on the pool wall in this version of the VMWM. It is likely that elements such as geometric cues, textures, and luminance variation in the distal region of the maze provided participants with information about the configuration of the environment that better enabled males than females to effectively orient themselves towards the correct quadrant. We cannot rule out, however, that participants also relied on subtle features in the proximal region such as light reflections, water flow, tiles at the bottom of the pool, or other elements within the confines of the pool. As to why such elements would be implicated in significantly different performance by sex, we cannot say. Some verisimilitude is lost when these proximal elements are not present, but future research should take them into consideration in the design of the VMWM.

Visible platform trials in the VMWM are implemented to ensure the participant understands the task and can complete it in spite of the virtual nature of the environment (Livingstone & Skelton, 2007; Newhouse et al., 2007; Sandstrom et al., 1998). In human research, normal performance in the VPTs is considered an indication that motivational, emotional, or sensorimotor defects are not a factor in completing the task (Livingstone & Skelton, 2007; Newhouse et al., 2007), and performance in the visible platform trials has been considered non-spatial. However, evidence in animal research has shown that subjects can still use a spatial strategy (D'Hooge & De Deyn, 2001). For example, Hauben, D'Hooge, Soetens, and De Deyn (1999) found that after completing the visible trials, mice showed preference for the platform quadrant in a subsequent probe trial, indicating "that the animals not only apply cue learning but also spatial search strategies, even when proximal cues are available. This spatial strategy enables them to show a preference for the target quadrant, and to enter the target area when neither the platform, nor the cue are present" (Hauben et al., 1999, p. 339). This suggests there might be more to performance in the visible trials than non-spatial considerations like lack of motivation or sensorimotor impairment.

In this study, participants completed the VPTs in approximately 10 s and the HPTs in 83 s, clearly demonstrating that neither motivation, sensorimotor impairments, nor difficulty navigating virtual environments were major factors in task performance. On the other hand, the results did yield a sex difference approaching significance, with males completing the task 1.7 s faster than females. Several studies have found similar results (see Sandstrom et al., 1998; Woolley et al., 2010). It is unknown why this difference should be found, though Woolley et al. (2010) suggest that sex differences in motivation, sensitivity to reward, interaction with virtual environments, and responsiveness to virtual stimuli may still be contributing factors. Sandstrom et al. (1998) found that females showed a slight hesitation upon first entering the maze, and some also mentioned feeling confused when they found themselves facing the room and not the pool. Given the ubiquity of movement restrictions imposed upon women across both time and space (Ardener, 1981) it is possible that these restrictions also influence women's navigational behaviour in virtual space (Murray, Chow, & Connolly, 2015). While a preliminary analysis of the SLs obtained in the VPTs showed that females hesitated at the beginning of the trial, the recorded latencies in the present study were not sensitive enough to detect movement at sufficiently fine-grained levels to conduct a thorough statistical analysis of the recorded movement data. Because in this study participants viewpoint at trial onset was random, sometimes facing the centre of the pool and sometimes facing away from it at different angles, the possibility of a spatial component at the beginning of each VPT cannot be discounted, as participants sought to orient themselves to get a better view of the pool and visually locate the platform. This may account for the results of our post hoc correlation analysis showing a statistically significant relationship between MRT scores and SLs in the VPTs: the higher the score on the MRT the faster the participant reached the visible platform. This significant correlation, further strengthened by the significant sex differences in MRT scores, suggests that a spatial component to the VPTs had a differential impact on males and females. Future research should measure and control the possible impact of extraneous variables to ensure that the VPTs truly represent a non-spatial component of the VMWM paradigm, and to identify any spatial processes, such as those activated by variation in starting position, that might contribute to search performance of the visible target.

The significant correlation between participants' MRT scores and maze performance we found is consistent with past research (Astur et al., 2004). We do not know why the ability to mentally rotate abstract geometrical 3D figures appears to be related to the ability to navigate real and virtual environments. The mental rotation task requires the visualization and angular transformation of discrete three-dimensional objects held in memory (Khooshabeh & Hegarty, 2010). Navigation, on the other hand, usually takes place in large-scale environments that surround the navigator's point of view and are encountered piecemeal. Researchers believe that to be useful as navigation aids, the environmental cues encountered along a route need to be configured allocentrically in relation to each other and egocentrically in relation to the navigator's point of view (Ward et al., 1986). Therefore, as a navigator moves through the environment, he/she develops an object-like 3D configuration, or mental map, of the environment. The VMWM in this experiment did not contain any distinct landmarks, so to the extent that participants developed a 3D configuration of the environment, they likely did so using configurational information such as room geometry obtained from the distal region of the room. If, as Dabbs et al. (1998) point out, “three-dimensional visualization promotes the use of abstract Euclidean navigation,” then mental rotation is likely to predict successful navigation in environments where the use of configurational information is necessary. The correlation between MRT scores and maze performance in the present experiment lends support to this hypothesis.

As predicted, we also found a significant sex difference in the MRT favouring males. However, even though we also found correlations in the expected direction between males' and females' MRT scores and maze performance, these correlations failed to reach statistical significance. Dividing the sample population into smaller groups likely resulted in a loss of power, and this is an issue that should be addressed in future research. As it stands, however, we did not find a significant difference between males and females in the extent to which their MR were associated with performance in the maze. Together with the finding that, at a general level, MR abilities predict performance in the maze, then individual differences in MR abilities are more important than the sex of the individual.

Although experts outperformed novices in mental rotation, videogame expertise did not predict performance in the VMWM task. Also, the interactions between sex and expertise were not significant, which suggests that experience with games did not account for the observed sex differences. The correlation analysis did show that novices' mental rotation scores correlated with performance in the hidden platform and probe trials. Uttal and Cohen (2012) suggest that as domain expertise increases, reliance on spatial abilities lessens. Thus, a greater reliance by novices on mental rotation to complete the maze could be explained by their lesser experience with videogames than experts. Generally, these results do not support the notion that videogame expertise is associated with superior spatial learning and memory in the absence of proximal cues, though they do show that gaming experience is associated with superior mental rotation and can be a factor in maze performance in cases where expertise is low.

7. The proximal cues experiment - results and discussion

In this experiment, we added five proximal cues around the pool wall of the VMWM. Following the work of Livingstone and Skelton (2007), we predicted that performance would improve for all participants, but of particular interest to us, following from the findings of Sandstrom et al. (1998), was examining the extent to which sex differences might be reduced with the availability of proximal landmarks. Past research has found that playing 3D action video games improves spatial abilities (Feng et al., 2007; Quaiser-Pohl et al., 2006), so we predicted that experienced videogame players would perform better than inexperienced players. A correlation analysis between maze performance and mental rotation was conducted to investigate whether the availability of proximal affected this relationship.

7.1. Results

7.1.1. Search latencies

7.1.1.1. Visible platform trials. The mean SLs and SEs for the VPTs are reported in Table 4.

The main effect of trial was not significant [$F(3, 99) = 2.17$, $p > .05$]. A significant interaction effect of trial with sex showed an unexpected difference between males and females in the VPTs [$F(3, 99) = 3.58$, $p < .05$], as illustrated in Fig. 14. The main effect sex approached significance [$F(1, 33) = 3.99$, $p = .05$] with males reaching the platform slightly faster than females.

The adjusted mean SLs by novices and experts (Fig. 15) show comparable performance between the two groups. The interaction effect of trial with expertise was not significant [$F(3, 99) = 1.97$, $p > .05$], and neither was the main effect of expertise [$F(1, 33) < 1$, $p > .05$].

Table 10 and Table 11 in Appendix C shows the adjusted mean SLs grouped by sex and expertise. These mean SLs are illustrated on Fig. 16 and Fig. 17. The interaction effect of trial with sex and expertise was not significant [$F(3, 99) = 1.42$, $p > .05$]. Further, the interaction effect of sex × expertise not significant [$F(1, 33) = 1.19$, $p > .05$].

7.1.1.2. Hidden platform trials. Table 5 shows the adjusted mean

Table 4
Adjusted mean search latencies (seconds) and standard errors (SE) in the visible platform trials.

Visible platform trial	Females (SE)	Males (SE)	Experts (SE)	Novices (SE)
1	9.9 (1.06)	12.4 (1.32)	9.6 (1.09)	12.7 (1.30)
2	11.5 (0.69)	8.6 (0.86)	10.7 (0.71)	9.5 (0.85)
3	11.2 (0.94)	8.3 (1.17)	9.9 (0.97)	9.5 (1.15)
4	9.7 (0.60)	7.7 (0.75)	8.6 (0.61)	8.8 (0.73)
Overall	10.6 (0.41)	9.3 (0.51)	9.7 (0.42)	10.1 (0.50)

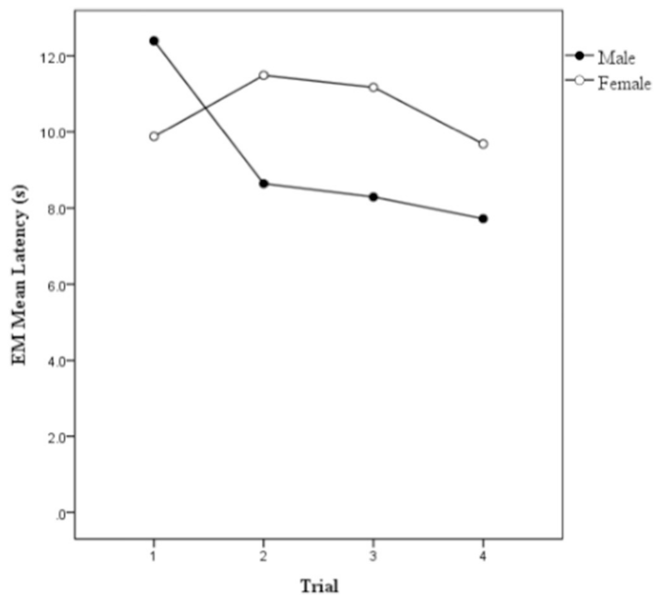

Fig. 14. Males' and females' adjusted mean search latencies (s) in the visible platform trials.

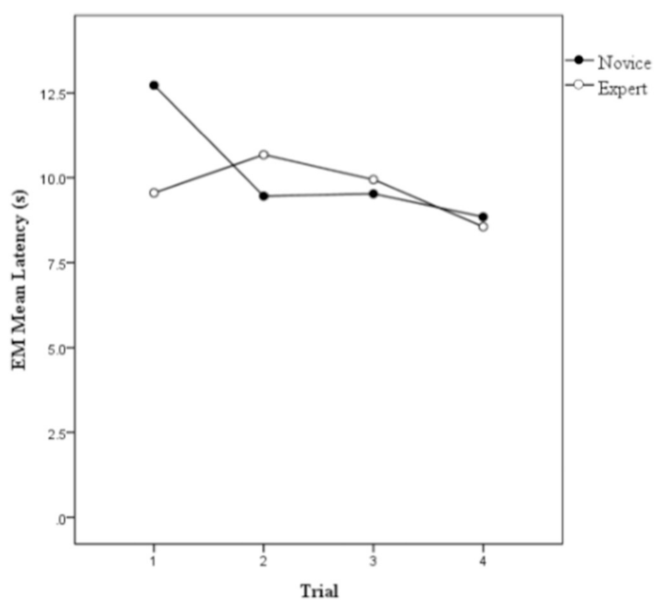

Fig. 15. VG experts' and novices' adjusted mean search latencies (s) in the visible platform trials.

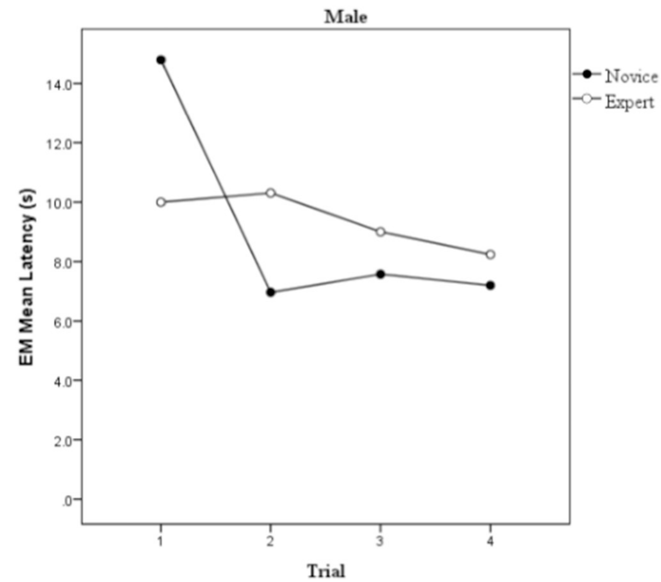

Fig. 16. Male VG experts' and novices' adjusted mean search latencies (s) in the visible platform trials.

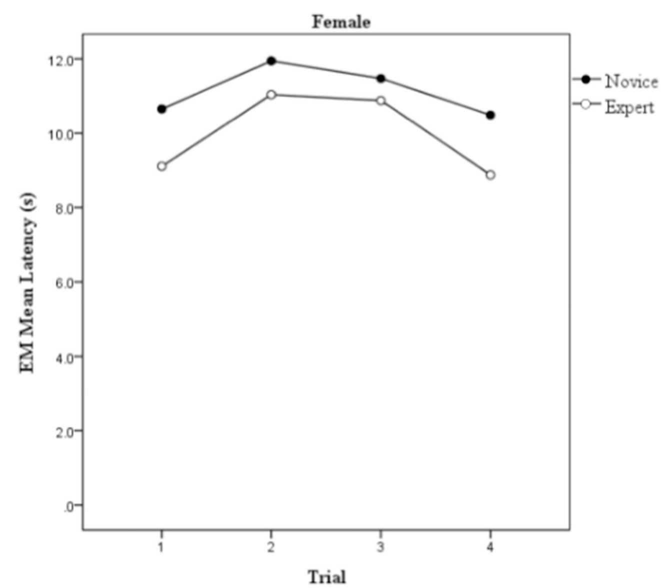

Fig. 17. Female VG experts' and novices' adjusted mean search latencies (s) in the visible platform trials.

latencies. There was a non-significant main effect of trial block [F(2, 66) = 1.29, $p < .05$] seen on Fig. 18, males obtained lower SLs than females, but the interaction effect of trial block and sex was not significant [F(2, 66) < 1, $p < .01$]. The analysis did yield a significant main effect of sex [F(1, 33) = 4.22, $p < .05$], which demonstrated that males obtained shorter SLs than females.

Although experts are shown to obtain lower latencies than novices across the three blocks of trials (Fig. 19), no significant interaction of trial block with expertise was found [F(2, 66) < 1.42, $p > .05$]. However, the main effect of expertise was significant [F(1, 33) = 7.73, $p = .01$].

Analyzing SLs based on sex × expertise (Table 12 and Table 13 in Appendix C) did not yield an interaction effect of trial block with sex and expertise [F(2, 66) < 1, $p > .05$]. The interaction effect of sex with expertise, on the other hand, was significant [F(1, 33) = 4.21, $p < .05$]. Post hoc LSD pairwise comparisons showed that female

Table 5
Adjusted mean search latencies (s) and standard errors (SE) in the HPT blocks.

HPT block	Females (SE)	Males (SE)	Experts (SE)	Novices (SE)
1	27.3 (4.092)	19.8 (5.10)	19.7 (4.20)	27.3 (5.01)
2	24.2 (2.683)	18.5 (3.35)	15.5 (2.76)	27.1 (3.29)
3	20.5 (1.639)	16.0 (2.04)	15.9 (1.68)	20.7 (2.01)
Overall	24.0 (1.804)	18.1 (2.25)	17.0 (1.85)	25.0 (2.21)

experts were faster than female novices ($p < .001$) but there was no significant difference between male experts and novices ($p > .05$). Male novices were faster than female novices ($p = .01$), but there were no significant differences between male and female experts ($p > .05$). The estimated marginal means grouped by sex and expertise are shown in Fig. 20 and Fig. 21.

7.1.1.3. Probe trial. In the probe trial, the main effect of sex was not

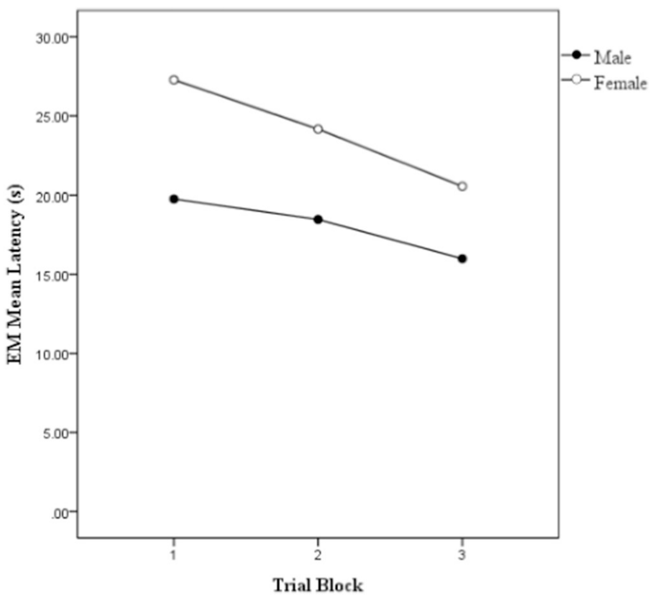

Fig. 18. Males' and females' mean search latencies (s) in the HPT blocks.

Table 11
Female VG experts' and novices' adjusted mean search latencies (seconds) and standard errors (SE) in the visible platform trials.

Trial	Experts (SE)	Novices (SE)
1	9.1 (1.73)	10.7 (1.22)
2	11.0 (1.13)	11.9 (0.80)
3	10.9 (1.54)	11.5 (1.09)
4	8.9 (0.98)	10.5 (0.69)
Overall	10.0 (0.66)	11.1 (0.47)

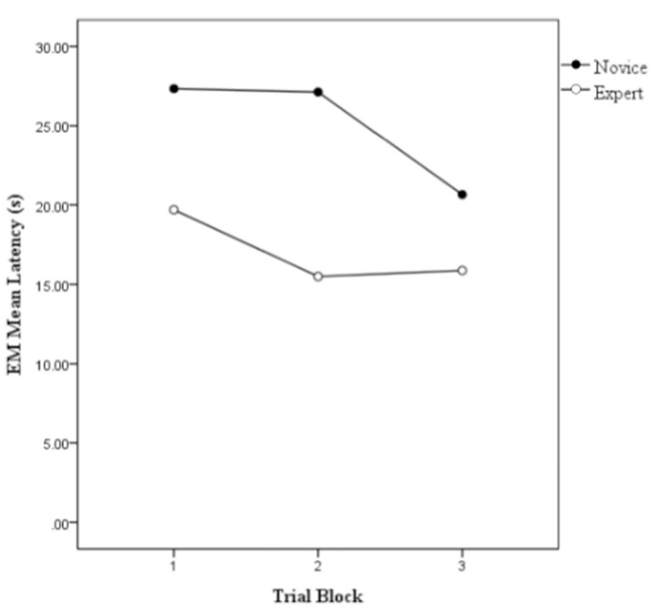

Fig. 19. VG experts' and novices' adjusted mean search latencies (s) in the HPT blocks.

statistically significant [$F(1, 33) < 1$, $p > .05$], nor was there a main effect of expertise [$F(1, 33) = 1.42$, $p > .05$]. The interaction effect of sex × expertise was not significant [$F(1, 33) = 1.75$, $p > .05$].

7.1.2. Mental rotation test - differences and correlations

7.1.2.1. Sex. The differences between males (13.4 ± 1.48 s) and females (10.8 ± 1.17 s) in MR scores were not statistically significant: $t(35) = 1.44$, $p > .05$, $r = 0.24$. We do not know why, for this experimental group, the differences were not as expected, although the differences that were demonstrated were in the expected direction (Fig. 22).

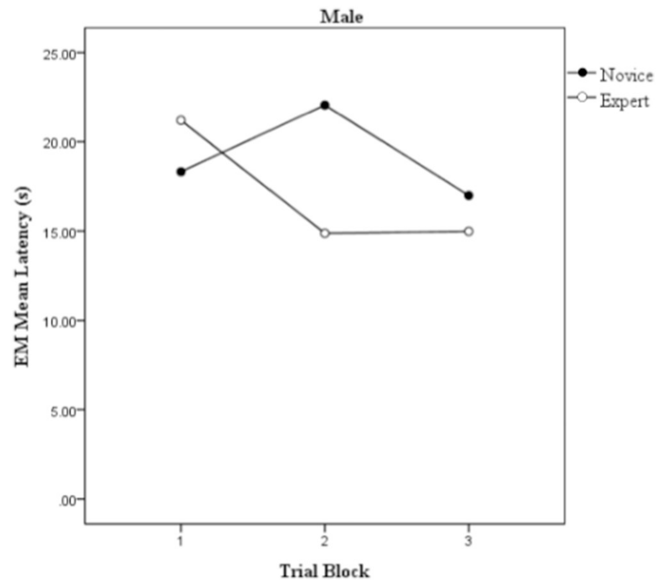

Fig. 20. Male VG experts' and novices' adjusted mean search latencies (s) in the hidden platform trials.

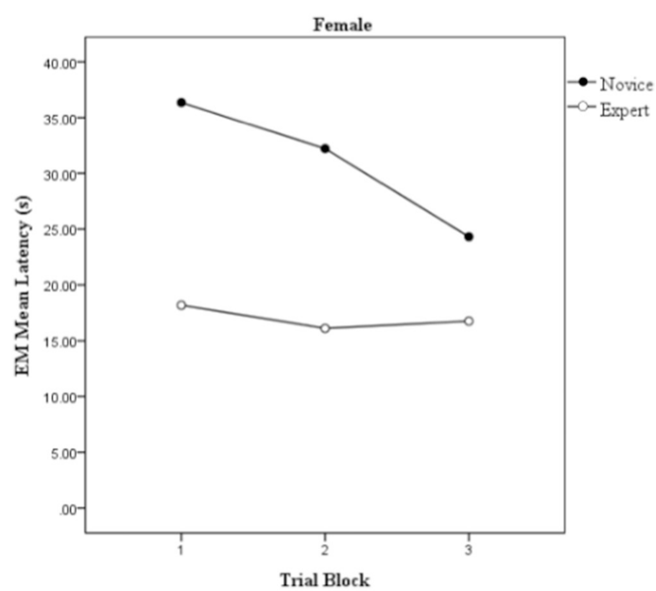

Fig. 21. Female VG experts' and novices' adjusted mean search latencies (s) in the hidden platform trials.

In general, MR scores did not correlate significantly with *SL* in the HP trials, $\tau = -0.13$, $p > .05$. There was no significant correlation between SL and females' MR scores, $\tau = -0.24$, $p > .05$, or males' MR scores, $r = 0.18$, $p > .05$. In the probe trial, MR scores were not significantly correlated with search dwell time on the platform quadrant, $\tau = 0.06$, $p > .05$. Separately, females' MR scores did not correlate significantly with quadrant dwell time, $\tau = -0.05$, $p > .05$, and neither did males' MR scores, $r = 0.12$, $p > .05$.

A post-hoc analysis of the correlation between MR scores and performance in the VPTs showed a general pattern in which the higher the MR scores the lower the SLs, but unlike the results in Experiment 1, this correlation failed to reach statistical significance, $r = -0.21$, $p > .05$.

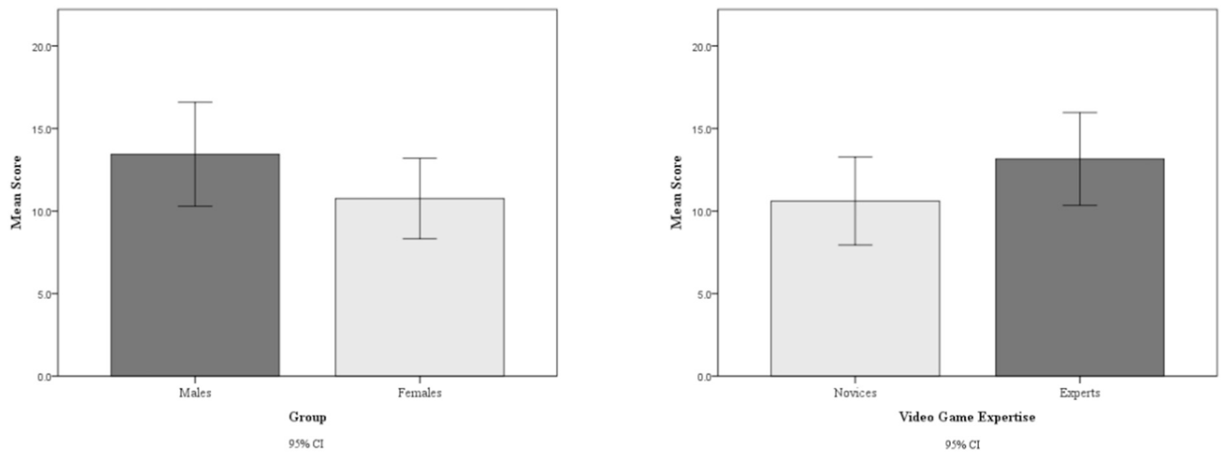

Fig. 22. Mental Rotation test scores grouped by sex and expertise separately.

7.1.2.2. Videogame expertise. In the mental rotation test (Fig. 22), the differences between novices (10.6 ± 1.26 s) and experts (13.2 ± 1.34 s) were statistically non-significant, $t(35) = 1.38$, $p > .05$, $r = 0.23$.

There were no significant correlations between novices' SL in the HP trials and their MR scores, $r = -0.31$, $p > .05$, or experts' MR scores, $r = -0.07$, $p > .05$. On the probe trial, there were no significant correlations between search time on the platform quadrant and novices' MR scores, $\tau = -0.06$, $p > .05$, or experts' MR scores, $r = -0.01$, $p > .05$. This result indicated that MR did not seem to play a significant role in spatial learning and memory when proximal cues were visible.

The mental rotation scores grouped by sex and expertise are shown on Fig. 23. While we cannot make statistical assertions in this regard, these results follow a consistent pattern favouring videogame experts across genders.

7.2. Discussion

In this second experiment, participants' search for the platform in the virtual maze was assisted by the provision of five landmark-type cues around the pool wall. On average, participants reached the platform in approximately 10 s when the platform was visible, a result which is comparable to that obtained in the first experiment. Thus, the addition of landmarks did not have a discernible effect on performance when the platform was clearly visible. Moreover, the low latencies in the VPTs indicate that in general participants understood the task, were motivated to complete it, and were not adversely affected by the virtual nature of the environment. It was unlikely that motor differences were a factor given that participants used the arrow keys to navigate the environment, and this did not require much physical strength or dexterity. However, just like in the first experiment, males reached the platform faster than females in the VPTs. Although the difference was small (~2 s), it approached significance in the first experiment and was significant in the second one. This is tentative evidence that the results obtained in the first study were not a matter of chance but reflect a genuine difference in performance. It is not clear why, even though the platform was visible, males and females performed differently. In the first experiment, we found that participants' MR scores significantly correlated with SLs in the VPTs, suggesting that the ability to perform spatial transformations came into play during the VPTs. As in experiment 1, viewing direction at trial onset was not fixed, raising the possibility that participants needed to first orient themselves to find the visible platform. However, although the pattern of correlation between MR and SLs was also negative in the VPTs of the Proximal Cues experiment, the correlation was not statistically significant. As well, the differences in MR between males and females in this second experiment were non-significant. So, in this experiment too, while we cannot entirely rule out MR as the reason why sex differences were found in the VPTs, other factors besides MR may have contributed to the sex difference in performance. As previously noted, prior research (Sandstrom et al., 1998) found females exhibited a short delay at trial onset in the VPTs and mentioned "confusion" because the view was that of the surrounding room instead of the pool, which may have created a further disadvantage to females when they first tried to orient themselves. In explaining the re-occurrence of gender differences among the second experiment's participants' times to reach a visible platform, we again note that women's restricted mobility throughout life across time and cultures (Ardener, 1981) may explain lower confidence and hesitation in initiating movement, but in that case it would be useful to assess the impact of these factors evident in the visible platform trials, and specifically whether they explain some part of the difference found in the hidden platform trials. However, this is not something we have seen taken into consideration by other researchers who have identified this difference in male and female performance in reaching a visible target. In our own study, the differences attributable to the delay were not large enough to explain, by themselves, the overall gender difference in performance in the hidden platform trials. Still we flag this observed initial discrepancy as potentially significant, and potentially

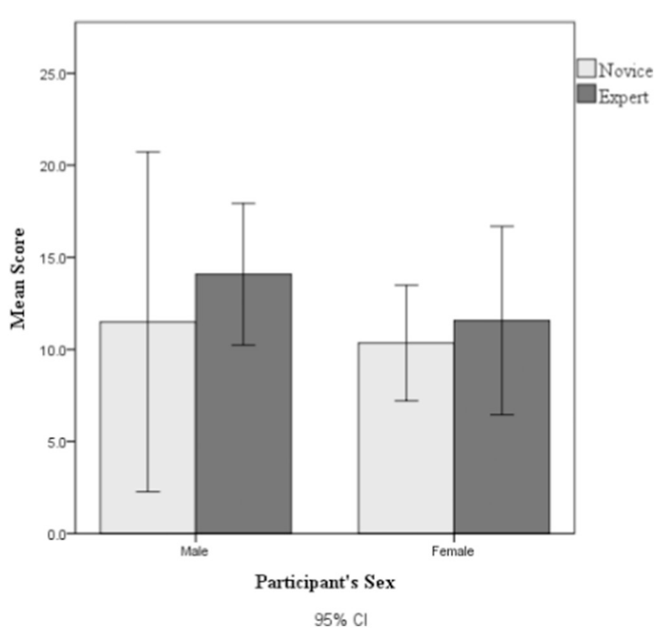

Fig. 23. Mental rotation scores grouped by sex × expertise.

compounded by other gender-specific contributing factors. Locking the point of view so that the participants face the pool at the beginning of trial may eliminate these differences or help determine whether non-spatial factors such as sex differences in motivation, confidence, sensitivity to reward, or familiarity with virtual environments in general exert an influence on performance.

As can be seen by the short latencies obtained in the HP trials, performance by all participants in the second experiment was much better than in the first experiment. The mean search time was 83 (SD = 40.16) seconds in the first experiment and approximately 23 (SD = 10.07) seconds in the second experiment. Therefore, the presence of proximal landmarks in the pool allowed all participants to more quickly orient themselves effectively and efficiently towards the platform. This improvement was also observed in the Probe trial. In the first experiment, participants spent 42% of their search time in the platform quadrant, compared to 77% in the second experiment. Participants in the second experiment mentioned that they used the visible landmarks to orient themselves and made no mention of other strategies beyond ensuring that their viewpoint was aligned with the landmark closest to the platform. It is unknown what specific spatial ability was associated with this search behaviour, though it is possible that the ability to visualize their own position in relation to the landmark improved participants' chances of finding the platform.

Egocentric perspective-taking is the “imagined movement of one's point of view in relation to other objects (or set of objects)” (Kozhevnikov, Motes, Rasch, & Blajenkova, 2006). The availability of landmarks in the present experiment likely promoted the use of perspective-taking over MR (though the two abilities may share some underlying components). In prior research, Kozhevnikov et al. (2006) found that mental rotation and perspective-taking predicted performance on route knowledge tasks, suggesting that both abilities share common spatial processes. However, the results of their research also showed that perspective-taking abilities were better predictors of landmark-pointing tasks and self-to-object environmental representations than MR (Kozhevnikov et al., 2006). Given that the search task in the second version of the maze required that the participants' update their position in relation to the proximal landmarks and the platform location, then it is feasible that they relied on perspective-taking more than on MR abilities. In contrast, the first version of the maze had no such landmarks, and therefore MR abilities would have been more useful in finding the platform in that environment. As Kozhevnikov et al. (2006) indicate, task-specific demands in egocentric and allocentric spatial transformations lead to the use of different abilities and strategies. Whereas the task demands in the first experiment promoted the use of an allocentric strategy because of the lack of visible proximal landmarks, the specific task demands of the second experiment may have promoted the use of an egocentric strategy. Future research should more closely examine these tentative findings by measuring not only MR but also abilities like perspective-taking, spatial visualization, and orientation. The VMWM provides a unique opportunity to disambiguate how different spatial abilities, along with other sensory, perceptual, and cognitive processes, influence spatial learning and memory.

While both males and females in the second experiment performed better than males and females in the first experiment, the improvement was particularly notable for females. The addition of landmarks greatly facilitated their learning and memory for the location of the platform. The sex differences in overall performance in the HPTs in the second experiment were small (about 9 s in favour of males) but, even so, were statistically significant. Sex differences in the probe trial disappeared altogether. The sex differences found in the second experiment demonstrate that learning the location of the hidden platform was not just a matter of randomly searching the general area near the key landmark (a crab image near the platform) but entailed refinement of the strategy across trials. To solve the search task, the participants were asked to look around while standing on the location of the platform. Given that the crab landmark was the closest landmark to the platform location, then participants looked at this landmark to encode the precise location of the platform. While all participants found the platform in a relatively short time, the fact that males were significantly faster than females, suggests that they were better able to map the location of the platform in relation to the proximal cues. If processes like perspective-taking, distance estimation, and spatial updating came into play, then the results further suggest a sex difference in these spatial abilities. This conjecture has found some support in the literature (see Chai & Jacobs, 2009; Lambrey & Berthoz, 2007). Of course, we cannot discount the possibility that other processes, such as personality factors (as per Coluccia & Louse, 2004) and spatial anxiety, influence performance. Indeed, Lawton (1994) and Lawton and Kallai (2002) found that females reported higher levels of spatial anxiety than males even across cultures and, in turn, spatial anxiety seemed to have inhibited their ability to orient themselves using environmental reference points. While the navigation task in our own experiment was not particularly anxiety-inducing as far as we could tell, aspects of the experimental conditions in the laboratory (e.g., male experimenter, unfamiliarity and starkness of room) could have been sources of anxiety. Minimizing and assessing spatial anxiety needs to be considered in future research that modifies conditions systematically to learn more about the extent to which anxiety affects performance. For now, this and other aspects of sex difference in spatial learning and memory remain unsolved. The design modifications made in the second experiment enabled us to demonstrate how sex differences can be reduced by controlling the type of information made visible to participants, a result that is consistent with the manipulation of distal landmarks by Sandstrom et al. (1998) in prior research. To explain the differences that remained, however, insofar as they may result from different cognitive processes, more research is needed.

While experts outperformed novices significantly in the HPTs when proximal cues were visible, there were no statistically significant differences in the VPTs and the Probe trial. Given that during the Probe trial spatial learning and memory was measured based on quadrant search time, these results indicated that the key difference between experts and novices lay in how accurate they were in learning the precise location of the small platform area during the HPTs. Although the adjusted mean difference was approximately 8 s in favour of experts in the HPTs, the difference was significant and consistent across trials. The significant interaction effect of sex and expertise showed that female experts were generally faster in reaching the platform. This pattern was not observed in males. Male novices did show faster overall latencies than female novices, but male experts did not significantly differ from female experts. Together, these results show that to the extent that videogame experience improves spatial learning and memory, females obtain more discernable benefits from this experience, a conclusion also arrived at in Feng et al. (2007).

The lack of significant correlation between novices' and experts' scores in the MRT and maze performance showed that in the presence of proximal cues, the small-scale ability to mentally rotate visualized objects was not significant in learning the location of the platform. This does not necessarily mean that the MR ability was not used at all, but to the extent that it was used, its influence on performance appears less salient than other spatial processes. Further, the non-significant differences between novices and experts in the MRT indicate that MR abilities alone cannot explain the differences between novices and experts found in the HPTs.

We previously discussed the possibility that abilities associated with *perspective-taking, spatial updating, and distance estimation* might play a role in learning the location of the platform. To demonstrate, the following illustration shows the search performance from a male participant during the HPTs of the PI experiment.

Fig. 24 shows a series of screenshots in a single HPT during which a male participant searched for the platform. The trial has been partitioned into several phases to better illustrate the sequence of events. After finding the platform, the avatar was held in place on top of the

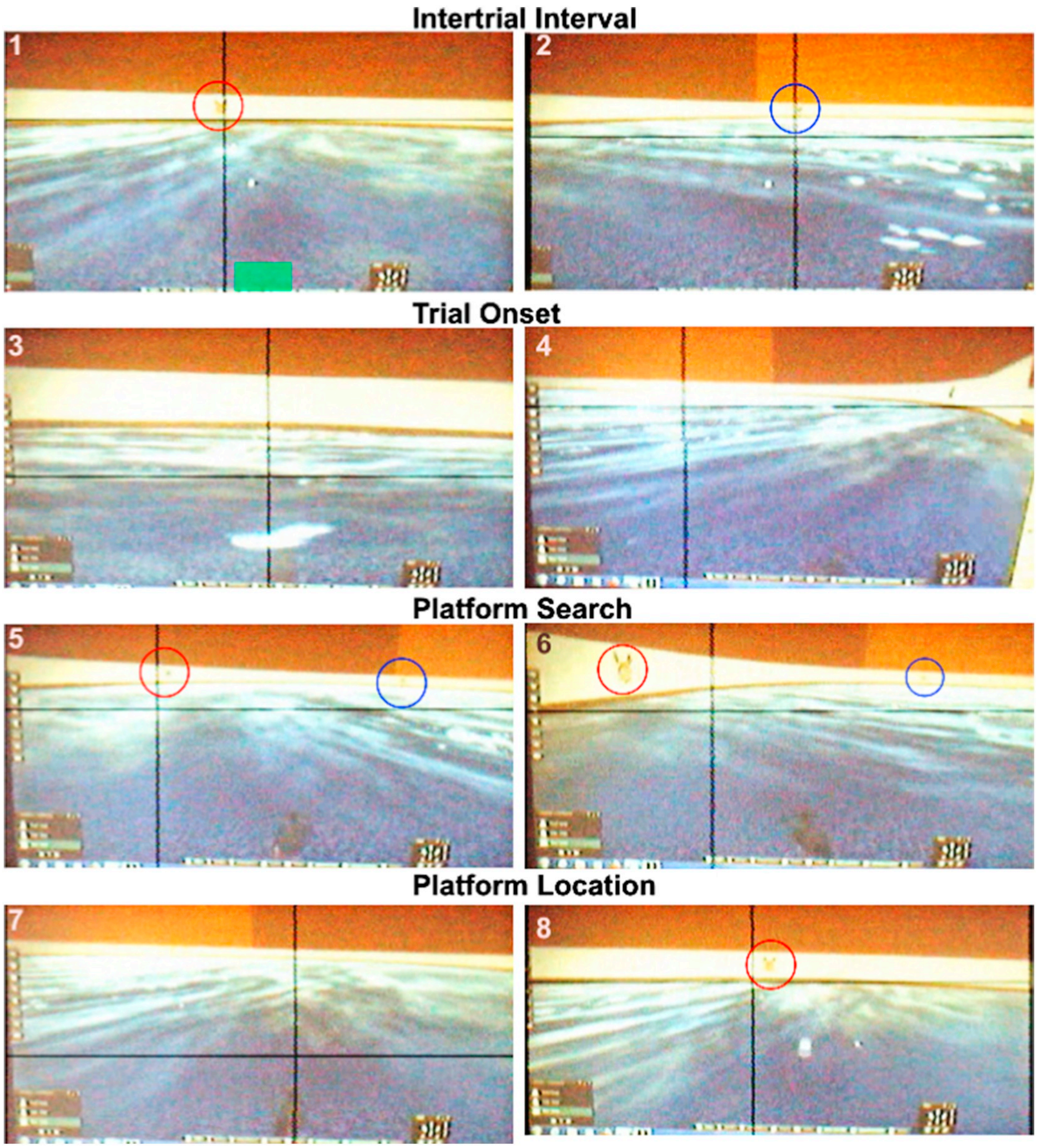

Fig. 24. Search performance by male participant in the HPTs of experiment 2.

platform and the participant was asked to look around and learn its location, which would remain in the same place for the duration of the remaining trials. In panel 1, the location of the rat avatar is shown by the green square shape. Panels 1–2 show this *intertrial interval* when the participant could only rotate to view the surrounding room. As panel 1 shows, the participant first gazed at the nearest proximal cue (a crab that has been circled in red for clarity), then he turned to the right to look at a second proximal cue farther away from the platform, a starfish along the pool wall (circled blue). We cannot say for certain what exact information (e.g., position relative to self, distance from platform, room geometry, luminance variation, and size of cue) was integrated at this point or whether the type of information extracted from the scene changed over the course of the trials. Once the participant was ready, he typed "OK" and teleported to start the next trial. At *trial onset* (panels 3–4), the participant began from one of four locations around the pool. In the shown example, he started the trial facing the pool wall (panel 3) then began to rotate the viewpoint towards the center of the pool (panel 4). He rotated in place and stopped only once he saw the crab cue (panel 5). This suggests that he was not just rotating so he could initiate movement forward but was actively searching for the proximal cue that he thought provided the best information in locating the platform. Once he located the crab cue (circled red), the participant moved towards it until he was within a short distance of it (panel 6). It is interesting to note that as shown in panel 6, the participant did not walk right up to the cue but stopped at a distance from it. Therefore, he had learned from previous experience that the platform location was at a specific distance from the pool wall and used memory of this fact to estimate where he needed to stop during the search trial. The second thing to notice in panel 6 is that the participant faced the crab cue at an angle that was not congruent with the perspective that he had observed during the intertrial interval when he was standing on the platform (as shown in panel 1). So, to better align his viewpoint to the viewpoint that he had learned during the intertrial interval, the participant proceeded to turn approximately 90 degrees clockwise away from the cue (panel 7) and moved forward until he was standing directly across from the cue, which was behind him and out of view. He then turned to face the cue (which had disappeared from his visual field), moved closer towards it, and successfully found the platform (panel 8). While all participants had unique ways of solving the task, this example shows that three processes of the search task were particularly important. First was the ability to quickly locate the most important cue in the environment and moved towards it. Second was the ability to recall and then estimate the distance between the cue (or pool wall) and the platform location. Third was the capacity to encode the perspective between the self and the cue as seen from the platform location and then bring this cognitive construct into alignment with the visual scene during the search task. While it is possible to rely on just one or two of these processes, using all three would likely produce a more efficient and effective performance. Moreover, although participants mentioned relying on the crab cue to orient themselves, it is possible that they integrated other information from other sources such as the geometric properties of the room. As Sandstrom et al. (1998) showed, participants capable of using different types of cues (landmark and geometric) are more likely to exhibit superior performance than participants who place greater reliance on a single type of cue. Thus, we cannot rule out the possibility that participants who showed superior performance in the present experiment did so because of their ability to integrate additional cues to orient themselves.

The preceding task analysis showed that the ability to learn and memorize directions, distances, and perspectives was essential for completing the VMWM task when proximal cues were visible. Novices and experts spent approximately 72% and 80% of their search time respectively on the correct quadrant during the Probe trial. Thus, both groups demonstrated spatial learning and memory. Given that the differences were not statistically significant, this result could be interpreted as evidence that when proximal cues are visible, videogame experience does not confer an advantage in solving the VMWM task. An alternate interpretation, however, is that the measure (dwell time in the platform quadrant) was not sensitive enough to detect subtle differences conferred by videogame experience. Such an interpretation is supported by the results obtained in the HPTs, in which experts outperformed novices by slight albeit statistically significant margin. The key difference between the two types of trials was the size of the area from which the measure was obtained. The area of measurement in the HPTs (i.e., the platform) was much smaller than that of the Probe trial (i.e., the quadrant), and thus the former required greater precision. Given that many videogames, and particularly 3D action ones, require high levels of precision to rapidly attend to and process transient events, it is evident how playing them can contribute to better performance on search tasks that require such levels of accuracy.

8. Conclusion

As we have noted, there is currently no standardized design for the VMWM. The focus in some studies, including the present one, has been the manipulation of landmark-type cues, which are visually salient. Less is known about the possible effect on spatial cognition of more subtle cues such as textures, geometrical shapes, luminance, colour, and other graphical and dynamic elements of the virtual maze. To more carefully investigate the factors that influence spatial cognition, as well as factors specific to virtual spatial tasks, more formalized versions of the maze will need to be designed and developed. The aim of a standard VMWM, or at minimum, of standard requirements to fully specify the design features of the environment being used, would be to minimize (or at least be enabled to take into account) differences across studies as much as possible, in order to improve the validity and reliability of the maze as a measure of spatial learning and memory.

The differences in performance found in both experiments have several implications for the use of virtual environments as tools for spatial research, assessment, and training. First, with respect to the VMWM specifically, is the importance of more precise descriptions of the version of the maze used in each experiment, since we have already discovered from our own research that maze design variations will matter decisively to shaping the results one arrives at, so there is work to be done before we can compare the results of experiments using variants of the maze.

Previous research has established that in processing spatial environments, both real-world and virtual, a variety of strategies and abilities can be called upon (e.g., Dabbs et al., 1998; Kallai, Makany, Karadi, & Jacobs, 2005; Sandstrom et al., 1998 and Ward et al., 1986) The results of our first experiment support the idea that there is a relationship between small-scale abilities such as MR and large-scale navigation. While we do not yet fully understand the relationship between navigation and small-scale spatial abilities, these two experiments provide some clues.

In the first, learning the location of the platform required the integration of distal information. A navigator could rely on MR to mentally manipulate a cognitive “map” of the environment to orient him/herself when landmarks are less useful for determining the location of a goal than other kinds of spatial cues. The results of the second study demonstrated the usefulness of proximal landmarks for accurately predicting the location of the platform. When proximal landmarks were available, MR did not appear to have played a significant role in spatial learning and memory. And though it's still possible participants integrated information from the distal region of the maze and directional information from other sources, our results made clear that they mainly relied on one specific landmark to find the platform. Questions remain about what additional spatial processes may have come into play during platform acquisition in the second study, but we suspect that perspective-taking and distance estimation played key roles. Coluccia and Louse (2004) have proposed that differences in visuospatial working memory may explain gender differences in spatially demanding tasks. The fact that gender differences in our study were greater when the task was difficult because of the lack of landmarks, lends some support to this notion. This topic has not been fully explored in the VMWM literature. The flexibility of the VMWM could, however, permit future research to more precisely delineate the role that MR and other spatial abilities play in place learning.

There were several limitations to the studies reported here. First, we did not have access to a large pool of participants, so an adequate balance between groups was not possible. This limited the statistical power of our results. Second, as is standard practice in the field, we relied on repeated measures analysis of variance to analyse the learning curves. As Young, Clark, Goffus, and Hoane (2009) point out, however, this approach lacks power and relies on post hoc tests that limit interpretation, and do not take into account the nonlinear trends that are sometimes found in the data. They offer an alternate approach that relies on regression analysis (i.e., mixed effects modelling) to address some of these limitations. Lastly, we based our measures of expertise on participant self-report, which can be unreliable. The nature of videogame expertise is not fully understood, and therefore measuring and assessing videogame experience is not an exact science. This, too, is an area that requires more precise instruments to accurately reflect the videogame experience and skills of the participants.

Research to date indicates that spatial abilities predict success in STEM fields, most particularly in the early stages of expertise development (Uttal & Cohen, 2012). This matters because of the centrality of STEM subjects to 21st century educational and vocational demands, the fact that STEM jobs pay more than non-STEM jobs, and because of the persistent under-involvement of girls and women in STEM learning and work. It is important, too, in view of evidence that “greater gender equality and economic development were associated with better performance” in spatial abilities like mental rotation (Lippa, Collaer, & Peters, 2010).

For educators seeking ways to assess students' spatial abilities, the VMWM provides a versatile environment. Through the alteration of design elements, the VMWM can be used to measure specific spatial abilities, as demonstrated by the results obtained when landmark visibility was manipulated. We learned from this study that such design changes could greatly diminish differences in performance between males and females, enabling females to achieve results comparable to males. However, in so far as different maze designs elicit reliance on different spatial skills to solve the search task, reducing performance differences by adding proximal cues requires us to identify more precisely what spatial skills are being called upon, and whether these skills can also support the kinds of spatial abilities associated with success in STEM fields. As Lawton (2010) states "There is a need for more research on ways to increase use of Euclidean navigational strategies, which, unlike route-based strategies, are useful even in environments that lack distinctive landmarks." (p. 331).

Interest in using videogames as tools for training to improve spatial cognition has increased in light of growing evidence that they do significantly improve certain kinds of spatial ability (Feng et al., 2007; Greenfield, 2009 and Oei & Patterson, 2013), though it is not yet understood what specific game elements are responsible for that improved performance. We explored the possibility that action VGs which feature 3D navigation may correlate with improvements in spatial learning and memory, perhaps mediated by corollary improvements in MR. While the VMWM task in this study did not possess the complexity found in modern VGs, some aspects of the virtual maze such as its three-dimensionality, and its avatar-driven navigation, did bear resemblance to such action-game environments. This allowed us to isolate and test specific elements that could be affecting spatial cognition. For instance, the availability of distal and proximal cues had a clearly different impact on how VG experts and novices performed in the search task, showing that differences between novices and experts were large when participants could not rely on proximal landmarks. Moreover, spatial performance in this context was correlated with MRT scores: higher MRT scores led to significantly faster platform acquisition when proximal landmarks were absent. The implication of this finding is that design of a virtual environment has a major impact on the type of cognitive abilities called upon to solve navigation tasks, suggesting that educators interested in using VGs to improve students' cognition need to carefully assess what cognitive abilities they wish to improve on, and selectively design environments to emphasize the use of those abilities. While there is not yet enough research to conclusively inform those decisions, it is expected that future research that deploys the affordances of virtual environments will be able to increasingly more systematically investigate specific relationships between videogame design and spatial cognition.

The growing body of evidence that demonstrates that spatial abilities are key factors in the early stages of STEM expertise, including the results of the present experiments, clearly shows a continuing disparity in these abilities between males and females. That disparity may also have implications for the differential representation of both sexes in an economy that emphasizes STEM expertise. To better understand the bases of these differences in spatial abilities and their impact on STEM representation, and to address persistent gender inequality in STEM subjects and fields, calls for a very targeted focus on creating conditions that allow and support women and girls to develop the abilities required for success in STEM occupations. Learning how to identify, evaluate and develop the skills and abilities associated with these cognitive domains is the goal of our experimental work using the VMWM to devise an easily administered valid and reliable test of spatial ability, and to discover which of these abilities might be developed through experience in the virtual space of videogame play.

Declaration of Competing Interest

None.

Acknowledgement

This research was supported by the Social Sciences and Humanities Research Council of Canada. Grant #435-2012-0586

Appendix A. Demographics

Table A.1
Demographic profile of participants who participated in the Distal Cues Experiment.

	English is second language - yes/no	Right/left handedness	Has taken courses in videogame design, 3D graphics and animation - yes/no	Involved in sports - yes/no
All (45)	27/18	42/3	11/34	38/7
Males (16)	8/8	14/2	6/10	15/1
Females (29)	19/10	28/1	5/24	23/6

Table A.2
Demographic profile of participants who participated in the proximal cues experiment.

	English is second language - yes/no	Right/left handedness	Has taken courses in videogame design, 3D graphics and animation - yes/no	Involved in sports - yes/no
All (37)	31/6	32/5	23/14	27/10
Males (16)	14/2	12/4	8/8	13/3
Females (21)	17/4	20/1	15/6	14/7

Appendix B. Videogame survey

B.1. Distal cues experiment

Responses to the videogame survey items were used to determine videogame expertise. Generally, most participants (~55%) stated that they frequently or often played videogames as children and only around 9% stated that they had never played videogames as children. Males (81%) were more likely than females (41%) to indicate that they played videogames frequently or often as children. At the time of the experiment, 31%, of participants reported being "active game players", with a higher proportion of males (56%) than females (17%) indicating being active game players. On average, participants played 7.3 h per week during the past year, with males (10.5 h) indicating that they engaged in videogame play almost twice as many hours as females (5.5 h).

It was important to find out how many participants had experience with precisely the kinds of games that previous research had found to improve the spatial abilities relevant to the present study. Most participants (42.2%) stated that they had never played a first-person shooter (FPS) game, while the rest indicated that they seldom (24.4%), sometimes (20.0%), frequently (8.9%), or often (4.4%) played such games. Grouped by sex, again we see a disparity showing that females (59%) were more likely than males (12.5%) to indicate that they had never played FPS games. At the other end of the scale, males (~31%) were more likely than females (~3%) to indicate that they played FPS games frequently or often.

To the question regarding what type of player the participants considered themselves to be, the majority (33.3%) thought of themselves as non-gamers and 6.7% viewed themselves as experts. The rest indicated that they were novice (17.8%), occasional (26.7), or frequent (15.6%) players. Consistent with other patterns, males (43.8%) were more like to state that they were frequent or expert players, while only 10.3% of females considered themselves frequent players and none consider themselves experts.

B.2. Proximal cues experiment

In response to how frequently they played videogames, the majority of participants (52%) indicated they frequently or often played videogames, while only 8% indicated playing never or seldom. Seventy-five percent of males stated that they played videogames frequently or often as children, and only 33% of females reported doing so. At the other extreme, only 6.3% of males mentioned never playing videogames as children but approximately 33% of females indicated that they never or seldom played videogames as children. Females (~33%) were more likely than males (~19%) to report only playing sometimes. Most participants (62%) did not consider themselves active players at the time of the experiment, and females (~67%) were more likely than males (~56%) not to consider themselves active players. On average, the number of hours that males (5.9 h) and females (5.4 h) engaged in videogame player per week during the year prior to the experiment was comparable.

The majority of participants (~70%) never or seldom played FPS games, 13.5% frequently or often did so, while the rest (~16%) indicated that they played action games sometimes. Twenty-five percent of males stated that they frequently or often played videogames, but only about 5% of females stated doing so. Most females (~95%) indicated that they never or seldom played FPS games, compared to 37.5% of males. A high percentage of participants (~40%) considered themselves non-gamers or novices, around 22% indicated being frequent or expert players, while approximately 38% stated that they were occasional players. More females (~52%) than males (25%) indicated being non-gamers or novices. Males (~38%) were more likely than females (~9%) to self-identify as frequent or expert game players.

Appendix C. Descriptive statistics

C.1. Distal cues experiment

C.1.1. Visible platform trials

Table 6
Male VG experts' and novices' adjusted mean search latencies (seconds) and standard errors (SE) in the visible platform trials.

Trial	Experts (SE)	Novices (SE)
1	9.7 (1.26)	12.0 (1.87)
2	8.7 (1.21)	9.6 (1.80)
3	7.8 (1.04)	10.2 (1.55)
4	8.3 (0.77)	7.5 (1.15)
Overall	8.6 (0.66)	9.847 (0.98)

Table 7
Female VG experts' and novices' adjusted mean search latencies (seconds) and standard errors (SE) in the visible platform trials.

Trial	Experts (SE)	Novices (SE)
1	9.9 (1.39)	12.6 (0.94)
2	10.4 (1.34)	10.8 (0.90)
3	10.3 (1.15)	10.5 (0.77)
4	9.1 (0.86)	10.6 (0.57)
Overall	9.95 (0.73)	11.1 (0.49)

C.1.2. Hidden platform trials

Table 8
Male VG experts' and novices' adjusted mean search latencies (seconds) and standard errors (SE) in the hidden platform trials.

Trial	Experts (SE)	Novices (SE)
1	82.6 (13.27)	103.9 (19.68)
2	41.7 (13.48)	74.2 (19.99)
3	36.9 (14.54)	50.6 (21.56)
Overall	53.7 (11.20)	76.3 (16.61)

Table 9
Female VG experts' and novices' adjusted mean search latencies (seconds) and standard errors (SE) in the hidden platform trials.

Trial	Experts (SE)	Novices (SE)
1	75.4 (14.67)	107.5 (9.84)
2	94.9 (14.90)	94.7 (10.0)
3	102.2 (16.07)	90.1 (10.78)
Overall	90.8 (12.38)	97.4 (8.31)

C.2. Proximal cues experiment

C.2.1. Visible platform trials

Table 10
Male VG experts' and novices' adjusted mean search latencies (seconds) and standard errors (SE) in the visible platform trials.

Trial	Experts (SE)	Novices (SE)
1	10.0 (1.32)	14.8 (2.29)
2	10.3 (0.86)	7.0 (1.49)
3	9.0 (1.17)	7.6 (2.03)
4	8.2 (0.75)	7.2 (1.29)
Overall	9.387 (0.51)	9.1 (0.88)

C.2.2. Hidden platform trials

Table 12
Male VG experts' and novices' adjusted mean search latencies (seconds) and standard errors (SE) in the hidden platform trials.

Trial	Experts (SE)	Novices (SE)
1	21.2 (5.10)	18.3 (8.84)
2	14.9 (3.35)	22.1 (5.80)
3	15.0 (2.04)	17.0 (3.54)
Overall	17.0 (2.25)	19.1 (3.90)

Table 13
Female VG experts' and novices' adjusted mean search latencies (seconds) and standard errors (SE) in the hidden platform trials.

Trial	Experts (SE)	Novices (SE)
1	18.2 (6.68)	36.4 (4.73)
2	16.1 (4.38)	32.2 (3.10)
3	16.8 (2.68)	24.3 (1.89)
Overall	17.0 (2.95)	31.0 (2.08)

References

Ardener, S. (1981). *Women and space: Ground rules and social maps.* Croom Helm, London: Bloomsbury Academic.

Astur, R. S., Ortiz, M. L., & Sutherland, R. J. (1998). A characterization of performance by men and women in a virtual Morris water task: A large and reliable sex difference. *Behavioural Brain Research, 93*(1–2), 185–190.

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*(1–2), 103–115.

Boot, W. R., Blakely, D. P., & Simons, D. J. (2011). Do action video games improve perception and cognition? *Frontiers in Psychology, 2*.

Boot, W. R., Kramer, A. F., Simons, D. J., Fabiani, M., & Gratton, G. (2008). The effects of video game playing on attention, memory, and executive control. *Acta Psychologica, 129*(3), 387–398.

Bowman, D. A., Kruijff, E., LaViola, J. J., & Poupyrev, I. (2004). *3D user interfaces: Theory and practice.* Redwood City, CA, USA: Addison Wesley Longman Publishing Co., Inc.

Brown, L. N., Lahar, C. J., & Mosley, J. L. (1998). Age and gender related differences in strategy use for route information. A map present direction giving paradigm. *Environment and Behavior, 30*(2), 123–143.

Bryant, K. J. (1982). Personality correlates of sense-of-direction and geographic orientation. *Journal of Personality & Social Psychology, 43*, 1318–1324.

Burnett, S. A., Lane, D. M., & Dratt, L. M. (1979). Spatial visualisation and sex differences in quantitative ability. *Intelligence, 3*, 345–354.

Carpenter, P. A., & Just, M. A. (1986). Spatial ability: An information processing approach to psychometrics. In R. J. Sternberg (Vol. Ed.), *Advances in the psychology of human intelligence. Vol. 3. Advances in the psychology of human intelligence* (pp. 221–253). Hillsdale, NJ: Erlbaum.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor analytic studies.* New York: Cambridge University Press Cambridge.

Chai, X. J., & Jacobs, L. F. (2009). Sex differences in directional cue use in a virtual landscape. *Behavioral Neuroscience, 123*(2), 276–283.

Clynes, T. (2016). How to raise a genius: Lessons from a 45-year study of supersmart children. *Nature, 537*, 154. Retrieved September 10, 2016, from http://www.scientificamerican.com/article/how-to-raise-a-genius-lessons-from-a-45-year-study-of-supersmart-children/.

Coluccia, E., & Louse, G. (2004). Gender differences in spatial orientation: A review. *Journal of Environmental Psychology, 24*(3), 329–340.

Dabbs, J. M., Chang, E.-L., Strong, R. A., & Milun, R. (1998). Spatial ability, navigation strategy, and geographic knowledge among men and women. *Evolution and Human Behavior, 19*(2), 89–98.

De Lisi, R., & Wolford, J. L. (2002). Improving children's mental rotation accuracy with computer game playing. *The Journal of Genetic Psychology: Research and Theory on Human Development, 163*(3), 272–282.

D'Hooge, R., & De Deyn, P. P. (2001). Applications of the Morris Water Maze in the study of learning and memory. *Brain Research Reviews, 36*(1), 60–90.

Driscoll, I., Hamilton, D. A., Yeo, R. A., Brooks, W. M., & Sutherland, R. J. (2005). Virtual navigation in humans: The impact of age, sex, and hormones on place learning. *Hormones and Behavior, 47*(3), 326–335.

Feng, J., Spence, I., & Pratt, J. (2007). Playing an action video game reduces gender differences in spatial cognition. *Psychological Science, 18*(10), 850–855.

Foreman, N. (1985). Algorithmic responding on the radial maze in rats does not always imply absence of spatial encoding. *Quarterly Journal of Experimental Psychology: Comparative and Physiological Psychology, 37B*(4), 333–358.

Ganis, G., & Kievit, R. (2015). A new set of three-dimensional shapes for investigating mental rotation processes: Validation data and stimulus set. Journal of Open Psychology Data, 3(1), e3.

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature, 423*(6939), 534–537.

Green, C. S., & Bavelier, D. (2007). Action video-game experience alters the spatial resolution of vision. *Psychological Science, 18*(1), 88–94.

Green, C. S., & Bavelier, D. (2012). Learning, attentional control, and action video games. *Current Biology, 22*(6), R197–R206.

Greenfield, P. M. (2009). Technology and informal education: What is taught, what is learned. *Science (New York, N.Y.), 323*(5910), 69–71.

Guizzo, F., Moe, A., Cadinu, M., & Bertolli, C. (2019). The role of implicit gender spatial stereotyping in mental rotation performance. *Acta Psychologica, 194*, 63–68.

Hauben, U., D'Hooge, R., Soetens, E., & De Deyn, P. P. (1999). Effects of oral administration of the competitive N-methyl-D-aspartate antagonist, CGP 40116, on passive avoidance, spatial learning, and neuromotor abilities in mice. *Brain Research Bulletin, 48*(3), 333–341.

Hegarty, M., Keehner, M., Cohen, C., Montello, D. R., & Lippa, Y. (2007). The role of spatial cognition in medicine: Applications for selecting and training professionals. In G. Allen (Ed.). *Applied spatial cognition*. Mahwah, NJ: Lawrence Erlbaum Associates.

Hegarty, M., Richardson, A. E., Montello, D. R., Lovelace, K., & Subbiah, I. (2002). Development of a self-report measure of environmental spatial ability. *Intelligence, 30*, 425–447.

Humphreys, L. G., Lubinski, D., & Yao, G. (1993). Utility of predicting group membership and the role of spatial visualization in becoming an engineer, physical scientist, or artist. *The Journal of Applied Psychology, 78*(2), 250–261.

Kallai, J., Makany, T., Karadi, K., & Jacobs, W. (2005). Spatial orientation strategies in Morris-type virtual water task for humans. *Behavioural Brain Research, 159*(2), 187–196.

Khooshabeh, P., & Hegarty, M. (2010). Representations of shape during mental rotation. *Cognitive shape processing. Symposium conducted at the AAAI Spring Symposia*Palo Alto, CA: AAAI March.

Kozhevnikov, M., Motes, M., & Hegarty, M. (2007). Spatial visualization in physics problem solving. *Cognitive Science, 31*(4), 549–579.

Kozhevnikov, M., Motes, M. A., Rasch, B., & Blajenkova, O. (2006). Perspective-taking vs. mental rotation transformations and how they predict spatial navigation performance. *Applied Cognitive Psychology, 20*(3), 397–417.

Lajoie, S. (2003). Individual differences in spatial ability: Developing technologies to increase strategy awareness and skills. *Educational Psychologist, 38*(2), 115–125.

Lambrey, S., & Berthoz, A. (2007). Gender differences in the use of external landmarks versus spatial representations updated by self-motion. *Journal of Integrative Neuroscience, 06*(03), 379–401.

Lawton, C. A. (1994). Gender differences in way-finding strategies: Relationship to spatial ability and spatial anxiety. *Sex Roles, 30*(11), 765–779.

Lawton, C. A. (2010). Gender, spatial abilities and wayfinding. In J. C. Chrisler, & D. R. McCreary (Eds.). *Handbook of research in psychology* (pp. 317–341). New York: Springer Press.

Lawton, C. A., & Kallai, J. (2002). Gender differences in wayfinding strategies and anxiety about wayfinding: A cross-cultural comparison. *Sex Roles, 47*(9), 389–401.

Linn, M. C., & Petersen, A. C. (1985). Emergence and characterisation of gender differences in spatial abilities: A meta-analysis. *Child Development, 56*, 1479–1498.

Lippa, R. A., Collaer, M. L., & Peters, M. (2010). Sex differences in mental rotation and line angle judgments are positively associated with gender equality and economic development across 53 nations. *Archives of Sexual Behavior, 39*(4), 990–997.

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*(1), 21–31.

Lohman, D. F. (1994). Spatial ability. In R. J. Sternberg (Vol. Ed.), *Encyclopedia of intelligence. Vol. 2. Encyclopedia of intelligence* (pp. 1000–1007). New York: Macmillan.

Lubinski, D., Benbow, C. P., & Kell, H. J. (2014). Life paths and accomplishments of mathematically precocious males and females four decades later. *Psychological Science, 25*(12), 2217–2232.

Malinowski, J. C. (2001). Mental rotation and real-world wayfinding. *Perceptual and Motor Skills, 92*(1), 19–30.

Margueles, J., & Gallistel, C. R. I. (1988). Heading in the rat: Determination by environmental shape. *Animal Learning & Behavior, 16*, 404–410.

McGuinness, D., & Sparks, J. (1983). Cognitive style and cognitive maps: Sex differences in representations of a familiar terrain. *Journal of Mental Imagery, 7*(2), 91–100.

Miller, L. K., & Santoni, V. (1986). Sex differences in spatial abilities: Strategic and experimental correlates. *Acta Psychologica, 62*, 225–235.

Moffat, S. D., Hampson, E., & Hatzipantelis, M. (1998). Navigation in a "virtual" maze: Sex differences and correlation with psychometric measures of spatial ability in humans. *Evolution and Human Behavior, 19*(2), 73–87.

Morris, R. (1984). Developments of a water-maze procedure for studying spatial learning in the rat. *Journal of Neuroscience Methods, 11*(1), 47–60.

Mueller, S. C., Jackson, C. P. T., & Skelton, R. W. (2008). Sex differences in a virtual water maze: An eye tracking and pupillometry study. *Behavioural Brain Research, 193*(2), 209–215.

Murray, J., Chow, E., & Connolly, C. (2015). Something in the way we move: Quantifying patterns of exploration in virtual spaces. *Proceedings of Foundation of Digital Games Conference* (Pacific Grove, CA).

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), 1–7.

Oei, A. C., & Patterson, M. D. (2013). Enhancing cognition with video games: A multiple game training study. *PLoS One, 8*(3), e58546. https://doi.org/10.1371/journal.pone.0058546.

Okagaki, L., & Frensch, P. A. (1994). Effects of video game playing on measures of spatial performance: Gender effects in late adolescence. *Journal of Applied Developmental Psychology, 15*(1), 33–58.

O'Laughlin, E. M., & Brubaker, B. S. (1998). Use of landmarks in cognitive mapping: Gender differences in self report versus performance. *Personality and Individual Differences, 24*(5), 595–601.

Orion, N., Ben-Chaim, D., & Kali, Y. (1997). Relationship between earth science education and spatial visualization. *Journal of Geoscience Education, 45*, 129–132.

Parsons, T. D., Larson, P., Kratz, K., Thiebaux, M., Bluestein, B., Buckwalter, J. G., & Rizzo, A. A. (2004). Sex differences in mental rotation and spatial rotation in a virtual environment. *Neuropsychologia, 42*(4), 555–562.

Pazzaglia, F., & Cornoldi, C. (1999). The role of distinct components of visual-spatial working memory in the processing of texts. *Memory, 7*(1), 19–41.

Peters, M., Laeng, B., Latham, K., & Jackson, M. (1995). A redrawn Vandenberg and Kuse mental rotations test: Different versions and factors that affect performance. *Brain and Cognition, 28*(1), 39–58.

Quaiser-Pohl, C., Geiser, C., & Lehmann, W. (2006). The relationship between computer-game preference, gender, and mental-rotation ability. *Personality and Individual Differences, 40*(3), 609–619.

Russell-Gebbett, J. (1985). Skills and strategies: pupils' approaches to three-dimensional problems in biology. *Journal of Biological Education, 19*(4), 293–298.

Sandstrom, N. J., Kaufman, J., & Huettel, S. A. (1998). Males and females use different distal cues in a virtual environment navigation task. *Cognitive Brain Research, 6*(4), 351–360.

Schmitz, S. (1997). Gender related strategies in environmental development: Effect of anxiety on wayfinding in and representation of a three-dimensional maze. *Journal of Environmental Psychology, 17*, 215–228.

Shepard, R. N., & Metzler, J. (1971). Mental rotation of three-dimensional objects. *Science, 171*(3972), 701–703.

Sims, V. K., & Mayer, R. E. (2002). Domain specificity of spatial expertise: The case of video game players. *Applied Cognitive Psychology, 16*(1), 97–115. https://doi.org/10.1002/acp.759.

Sneider, J. T., Hamilton, D. A., Cohen-Gilbert, J. E., Crowley, D. J., Rosso, I. M., & Silveri, M. M. (2015). Sex differences in spatial navigation and perception in human adolescents and emerging adults. *Behavioural Processes, 111*, 42–50.

Sorby, S. (2009). Developing spatial cognitive skills among middle school students. *Cognitive Processing, 10*, 312–315 Suppl. 2.

Sorby, S., & Baartmans, B. (1996). A course for the development of 3-D spatial visualization skills. *Engineering Design Graphics Journal, 60*(1), 13e20.

Sorby, S., & Baartmans, B. (2000). The development and assessment of a course for enhancing the 3-D spatial visualization skills of first-year engineering students. *Journal of Engineering Education,* 301e307.

Stieff, M. (2007). Mental rotation and diagrammatic reasoning in science. *Learning and Instruction, 17*, 219–234.

Suzuki, S., Augerinos, G., & Black, A. H. (1980). Stimulus control of spatial behavior on the eight-arm maze in rats. *Learning and Motivation, 11*, 1–18.

Talbot, K. F., & Haude, R. H. (1993). The relation between sign language skill and spatial visualization ability: Mental rotation of three-dimensional objects. *Perceptual and Motor Skills, 77*(3_suppl), 1387–1391.

Tarr, M. J. (1995). Rotating objects to recognize them: A case study on the role of viewpoint dependency in the recognition of three-dimensional objects. *Psychonomic Bulletin & Review, 2*(1), 55–82.

Uttal, D. H., & Cohen, C. A. (2012). *Spatial thinking and STEM education: When, why, and how?* ResearchGate*Vol. 1*.

Vandenberg, S. G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills, 47*(2), 599–604.

Voyer, D., Voyer, S., & Bryden, M. P. (1995). Magnitude of sex differences in spatial abilities: A meta-analysis and consideration of critical variables. *Psychological Bulletin, 117*(2), 250–270.

Wai, J., Lubinski, D., & Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology, 101*(4), 817–835.

Waller, D. (2000). Individual differences in spatial learning form computer-simulated environments. *Journal of Experimental Psychology: Applied, 6*, 307–321.

Ward, S. L., Newcombe, N., & Overton, W. F. (1986). Turn left at the church, or three miles north: A study of direction giving and sex differences. *Environment and Behavior, 18*(2), 192–213.

Williams, C. L., Barnett, A. M., & Meck, W. H. (1990). Organizational effects of early gonadal secretions on sexual differentiation in spatial memory. *Behavioral Neuroscience, 104*(1), 84–97.

Williams, C. L., & Meck, W. H. (1991). The organizational effects of gonadal steroids on sexually dimorphic spatial ability. *Psychoneuroendocrinology, 16*(1–3), 155–176.

Woolley, D. G., Vermaercke, B., de Beeck, H. O., Wagemans, J., Gantois, I., D'Hooge, R., & Wenderoth, N. (2010). Sex differences in human virtual water maze performance: Novel measures reveal the relative contribution of directional responding and spatial knowledge. *Behavioural Brain Research, 208*(2), 408–414.

Young, M. E., Clark, M. H., Goffus, A., & Hoane, M. R. (2009). Mixed effects modeling of Morris water maze data: Advantages and cautionary notes. *Learning and Motivation, 40*(2), 160–177. https://doi.org/10.1016/j.lmot.2008.10.004.