

## Research report

Sex differences and correlations in a virtual Morris water task,  
a virtual radial arm maze, and mental rotation<sup>☆</sup>Robert S. Astur<sup>a,\*</sup>, Jennifer Tropp<sup>b</sup>, Simona Sava<sup>b</sup>,  
R. Todd Constable<sup>a,c</sup>, Etan J. Markus<sup>b</sup><sup>a</sup> Department of Diagnostic Radiology, Yale University, New Haven, CT 06520, USA<sup>b</sup> Department of Psychology, University of Connecticut, Storrs, CT 06269, USA<sup>c</sup> Department of Neurosurgery, Yale University, New Haven, CT 06520, USA

Received 16 February 2002; received in revised form 3 February 2003; accepted 18 August 2003

## Abstract

Different tasks are often used to assess spatial memory in humans compared to nonhumans. In order to bridge this paradigmatic gap, we used a within-subject design to test 61 undergraduates on three spatial memory tasks. One of these tasks, the Vandenberg 3D mental rotation task, is classically used to assess spatial memory in humans. The other two tests are virtual analogues of two tasks used classically to assess spatial memory in rodents: the Morris water task and an eight-arm radial maze. We find that males perform significantly better than females on the mental rotation task and in finding a hidden platform in the virtual Morris water task. Moreover, during a probe trial, males spend significantly more distance of their swim in the training quadrant, but males and females do not differ in navigating to a visible platform. However, for the virtual eight-arm radial maze, there is no sex difference in working memory errors, reference memory errors, or distance to find the rewards. Surprisingly, an examination of the correlations among the three tasks indicates that only mental rotation ability and Morris water task probe trial performance correlate significantly among the three tasks (i.e. there are no significant correlations with traditional measures the tasks, e.g. time or distance to completion). Hence, the Morris water task and the eight-arm radial maze do not assess spatial memory in the same manner, and even after equating factors such as motivation, stress, and motor demands, there still are procedural demands of the tasks that reinforce differential strategy selection during spatial memory. This suggests that caution should be taken when utilizing these two tasks interchangeable as tests of spatial memory.

© 2003 Elsevier B.V. All rights reserved.

**Keywords:** Radial arm maze; Morris water task; Hippocampus; Virtual reality; Memory; Sex differences; Spatial navigation; Mental rotation

The hippocampus (HPC) has long been implicated as being critical for creating and maintaining spatial/cognitive maps [24]. For example, both rats and pigeons with HPC damage display spatial navigation impairments [5,21,34]. Physiologically, cells in the HPC that display increased firing rates when the animal is in a certain spatial location (i.e. place cells) have been a foundation of its involvement in spatial memory [23]. Ethologically, it has been shown both in birds [16] and rodents [14] that hippocampal size is significantly larger in species whose behavior depends on spatial ability compared to similar species with less of a dependence on spatial skills.

Initial experiments of spatial memory were conducted by Tolman [35] using rodents, since this type of memory was particularly naturalistic and easy to learn for rodents. This spawned a rich number of paradigms based on the premise of having an animal either learn to find a goal object or to avoid an aversive location. Resulting from this approach were mazes such as a starburst maze, spiral Battig maze, radial mazes, open-area mazes, runway mazes, and water mazes. Each maze type has made various unique contributions to the understanding of spatial memory. However, currently there are two mainstays of mazes that are commonly used to assess spatial learning and memory: the radial arm mazes and the water maze.<sup>1</sup>

<sup>☆</sup> For information on obtaining the virtual Morris water task or virtual radial arm maze see <http://www.nrc-iol.org/personnel/astur/astur.htm>.

\* Corresponding author. Present address: Olin Neuropsychiatry Research Center, Institute of Living, 200 Retreat Avenue, Hartford, CT 06106, USA. Tel.: +1-860-545-7776; fax: +1-860-545-7510.

E-mail address: [Robert.astur@yale.edu](mailto:Robert.astur@yale.edu) (R.S. Astur).

<sup>1</sup> Technically, it is not a “maze,” since a maze must, by definition, include a network of passages. Nonetheless, the term “water task” and “water maze” are often used interchangeably in the literature, and this liberty will also be taken here.

Radial arm mazes (RAM) typically involve a center area with a number of identical arms radiating outwards [26]. Each arm has a well at the distal end that may be baited with food. In a standard eight-arm version, four of the arms have food at the end and four arms do not. During a trial, an animal must retrieve all four rewards, after which the animal is removed from the maze. For subsequent trials, the same four arms are rewarded. With training, rodents learn to retrieve all four rewards without venturing into the never-rewarded arms. Moreover, after retrieving a reward from one of the arms, the animal remembers not to re-enter that arm again during the same trial. Hence, the measure of whether the animal remembers which arms are always baited (alternatively, which arms are never baited) is commonly used as a measure of reference memory. The extent to which the animal remembers which arms it has previously visited during a single trial is commonly used as a measure of working memory.

The water maze involves placing an animal in a round pool of cool, opaque water in which one area of the pool contains a fixed goal escape platform submerged a few centimeters under the surface of the water [20]. Whereas rodents are excellent swimmers, they find the cool water aversive, and they swim in search of an escape. Initially, the rodent searches around the pool walls for an escape and gradually ventures off from the wall and finds the platform. With training, normal rats eventually learn to use the distal cues in the room to navigate to the platform in a straight line regardless of their start position.

Commonly, these two tasks have been used interchangeably as tests of spatial memory and, because they are sensitive to HPC damage, also of HPC functioning [21,25,27,34]. Moreover, these tasks have been used to examine sex differences in rodents, and have revealed that male rats tend to perform better than females on a RAM, particularly early in training [30,38]. However, this sex difference is less evident when using the water maze [28].

As has been thoroughly pointed out by Hodges [12], there are a number of obvious differences between these two tasks which clouds interpretation of the behavioral discrepancies between the two tasks. Notably is the fact that the RAM uses appetitive motivation, whereas the water maze uses aversive motivation. Because of this, the water maze often has a higher level of stress than the RAM, and swimming in the water maze is more physically demanding than walking in the RAM. Moreover, the number of possible routes is more restricted in the RAM whereas there are fewer route limitations in water mazes. Also, the RAM contains a working memory component whereas the typical water maze water maze does not (although task parameters may be modified to test working memory using the water maze). Accordingly, there are reported discrepancies in spatial memory performance following HPC damage depending on whether the RAM or water maze is utilized to assess spatial memory. For example, Nunn and Hodges [22] report ischemia-induced deficits in the water maze but not in the RAM. However, the opposite result is reported by Davis and Volpe [10]:

ischemia-induced deficits are observed in the RAM, but not in the water maze. Hence, it is unclear what factors are contributing to these reported deficits in these tasks.

With the rise in computer technology, many of these paradigms have been adapted to be used with humans. This is advantageous in that spatial memory now can be tested in humans in a manner that is similar to how they are tested with rodents. Specifically, before this advancement, many of the classic spatial memory tests in humans assessed egocentric memory such as examined in mental rotation [7,36], finger mazes [8], and object location [33]. Alternatively, for nonhumans, allocentric memory is being assessed in tasks such as the water maze and the RAM. Some attempts have been made to test allocentric memory in humans in actual mazes [6,37], but such attempts are often cumbersome, require large spaces, and are taxing to the subjects. Moreover, these life-size mazes are not able to be adapted for functional imaging studies.

The use of virtual environments has proven to be successful for examining allocentric spatial memory in humans. For example, virtual environments have been used to examine basic learning phenomenon [13], sex differences [2,11,31] as well as hippocampus functioning [1,4,18]. Collectively, these experiments parallel well with the nonhuman research, specifically converging on the theme that the hippocampus is both involved and necessary for navigating through virtual environments [3,4], as well as documenting that males and females utilize different strategies to navigate through their environments. For example, Astur et al. [2] have shown repeatedly that males consistently navigate more efficiently in a virtual Morris water task, and Sandstrom et al. [31] have dissected this sex difference to reveal that females often rely on landmark information to navigate in a virtual environment whereas the males use both the landmark and the geometry information to navigate. Lastly, given that these mazes consist of virtual environments, the experimenter has complete control over all cues, movement options, and other relevant parameters to allow for more precise experimental control and manipulation. Hence, these virtual mazes seem to have many advantages over traditional spatial memory paradigms in terms of their flexibility, generalizability, and experimental control.

As mentioned previously, the discrepancies of performance impairments between the RAM and the water maze often are confounded by the fundamental differences in task demands such as the stress factors and intensive motor demands evident in the water maze but absent in the RAM. However, these confounds are eliminated in virtual versions of these tasks. Hence, by employing virtual versions of these tasks, we are in an excellent position to examine how these two classic tasks of spatial memory differ after equating levels of stress and motor demands.

Given that there have been robust reports of easily replicable sex differences in the virtual water maze [2,15,31] it is logical to assume that the virtual RAM would also reveal robust sex differences if it assessed spatial memory in the

same manner as the virtual water maze. To examine the validity of this assumption, we used a within-subject design to test undergraduates on both virtual RAM and a virtual water maze. In addition, all subjects were also tested on a mental rotation task because this task is a common test of spatial memory that is not dependent upon allocentric memory, and it is known to elicit robust differences between males and females. Lastly, given that this is a within-subject design, we correlated performance across these tasks to examine the relationships between these three tests of spatial memory.

## 1. Method

### 1.1. Participants

Sixty-one undergraduates (37 females; 24 males; 19.4 average (4.8 S.D.) years) from the University of Connecticut were recruited from Introductory Psychology classes. All participants had normal or corrected-to-normal vision and received class credit for their participation. Approval for this study was obtained from the University of Connecticut Institutional Review Board. Three additional subjects (all fe-

males) terminated the study early due to motion sickness from the virtual tasks.

### 1.2. Apparatus

An IBM-compatible computer with a SVGA color monitor was used for testing. Participants navigated through the mazes by manipulating a joystick. A speaker connected to the computer was used to provide auditory feedback to the participants.

### 1.3. Procedure

After obtaining informed consent, participants were given both verbal and written instructions on how each task would proceed. Before completing the two virtual mazes, all participants were tested on the mental rotation task. Afterward, half the participants started with the RAM followed by the water maze, and vice versa for the other participants. All participants were told to use the joystick to navigate. The view on the screen was a first-person view so that if they pushed the joystick to the right, the view on the screen would pan to the right and so on with other joystick movements (see Fig. 1). Both virtual rooms had

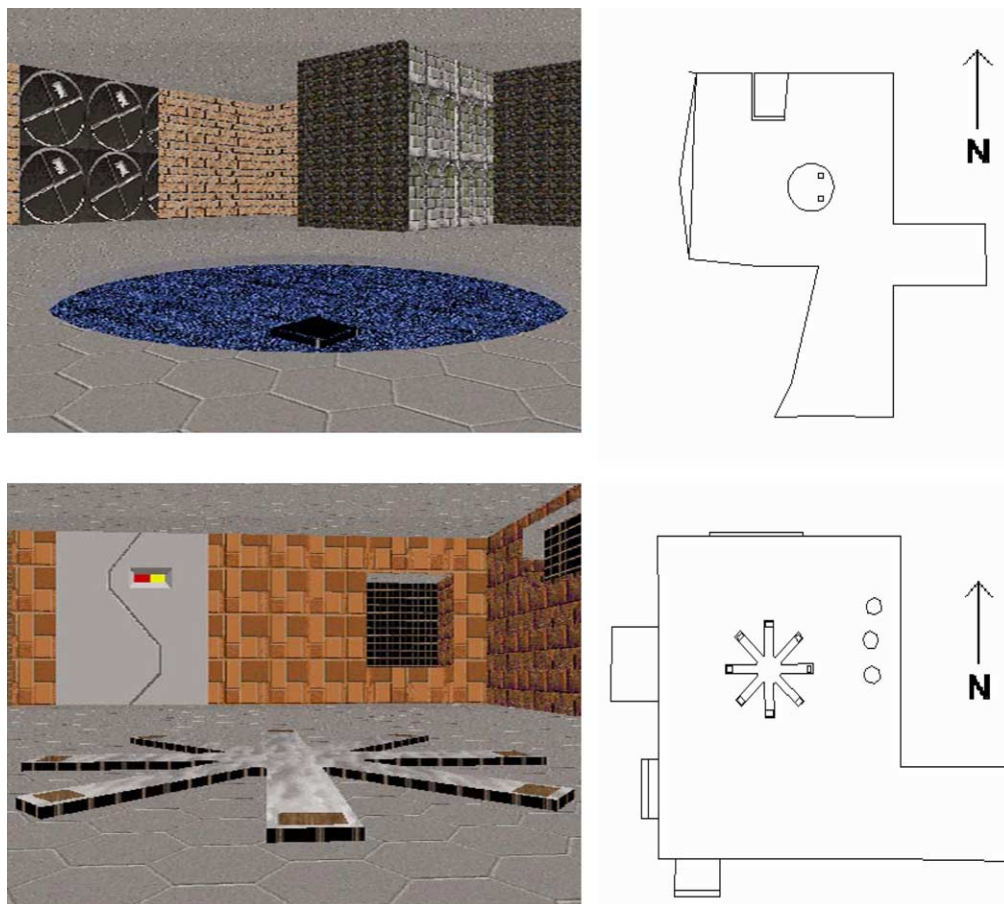


Fig. 1. Extra-maze views of the virtual tasks. Top: picture within (left) and aerial view of the room (right) for the virtual Morris water task. Bottom: picture within (left) and aerial view of the room (right) for the virtual radial arm maze. All views were in color.

a variety of doors, windows, and landmarks. Also, both rooms had various adjoining rooms and geometry to make the rooms nonsymmetrical. Participant location was written to a computer data file at approximately 10 Hz during each trial. After completing all the tasks, participants answered a questionnaire that queried their computer game experience as well as head trauma/neurological history. Individuals with a history of head trauma or neurological disorders were excluded from the study. For the females, additional information was obtained regarding where they were in their menstrual cycle and whether they were on any birth control that utilizes artificial hormones. However, this data was not used because of the 10 women who were menstruating, 3 were on artificial hormone birth control, and the other 7 were varied on where they were in their menstruation phase (e.g. some were on day 1 of menstruation and others were on day 5) so that the statistical power was too low to make any meaningful comparisons. We currently have ongoing studies examining these factors with a larger sample size.

#### 1.3.1. Mental rotation task

A pen and paper mental rotation task adapted from Vandenberg and Kuse [36] was used. Specifically, participants were given a target object and from four choices, they were to pick the two choices that matched the target object but which were simply rotated from the target. The other two choices did not match. Subjects were given 4 min to complete 12 of these problems. One point was given for each correct response, and participants were instructed to work as quickly as possible without compromising accuracy. The highest possible score was 24 correct (2 per problem). This task reveals strong sex differences in humans [7].

#### 1.3.2. Radial arm maze

Participants were instructed that they would find themselves in a virtual room that had eight runways extending out of a round middle area. They were further instructed that at the end of each runway is a well and that four of the runways have rewards in their well, and four of the runways do not. They were told to retrieve the four rewards as quickly as possible. Upon discovering a reward, a pleasant tone sounded. After discovering all four rewards, a pleasant chord played, and the following congratulatory text message was displayed: "Congratulations. You have found all the rewards." The screen then blanked, and participants were then teleported back to the middle area and were to begin the next trial after an ITI of 5 s. If 3 min elapsed and the four rewards were not found, the trial was terminated. Twenty trials were administered. The same configuration of rewarded arms (NW, W, S, and NE) was used for all subjects. A reference memory error was scored if the participant entered into an arm that was never-rewarded. A working memory error was scored if the participant entered into an arm that they had previously entered during that trial regardless of whether that arm was rewarded. So, for example, if the participant re-entered an arm that was never-rewarded,

that would be scored as both a reference memory error and a working memory error.

#### 1.3.3. Water maze

Participants were instructed that they will find themselves in a virtual pool, and that their goal would be to escape from the water as quickly as possible by swimming to a hidden platform. Procedurally, participants started from four different locations (north, south, east, and west) five times each for a total of 20 trials with a 5 s ITI. If the participant swam over the area of the pool where the platform is located, a tone sounded, the platform rose slightly out of the water, and a message saying "Congratulations. You have found the platform" was displayed. At this point, the participants were allowed free swimming movement for 3 s after which the trial terminated. This phase is referred to as "hidden platform" training.

After the 20 training trials, a probe trial was given in which the platform was removed from the pool, and the participant was allowed to search for the platform for 30 s, after which the trial terminated. There was no indication to the participant that the probe trial was in any way different than the previous 20 training trials, until the probe trial was already completed.

After the probe trial, the platform was moved to a different location in the pool and was raised slightly out of the water so that it was visible to the participant. Participants started from four different locations two times each for a total of eight trials. This phase is referred to as "visible platform" training. All events and consequences were identical to those in the hidden platform phase of testing.

Upon completion of the computerized testing, participants were given a questionnaire that queried their age, sex, and experience playing 3D computer games. Specifically, participants marked on a 4-point Likert scale the frequency with which they played such games (0, never; 1, rarely; 2, occasionally; 3, frequently). In addition, participants were also asked a brief question about whether they had any neurological problems or had any type of brain imaging at any point in their life. The duration of the experiment was approximately 60 min.

## 2. Results

### 2.1. Mental rotation

For the Vandenberg mental rotation task, males are significantly better at mental rotation than females,  $t(54) = 3.39$ ,  $P < 0.001$  (Fig. 2).

### 2.2. Morris water maze

All ANOVAs are repeated measures with Trial as the repeated measure. There is a significant effect of Trial, indicating that participants traveled less distance,  $F(19, 969) =$

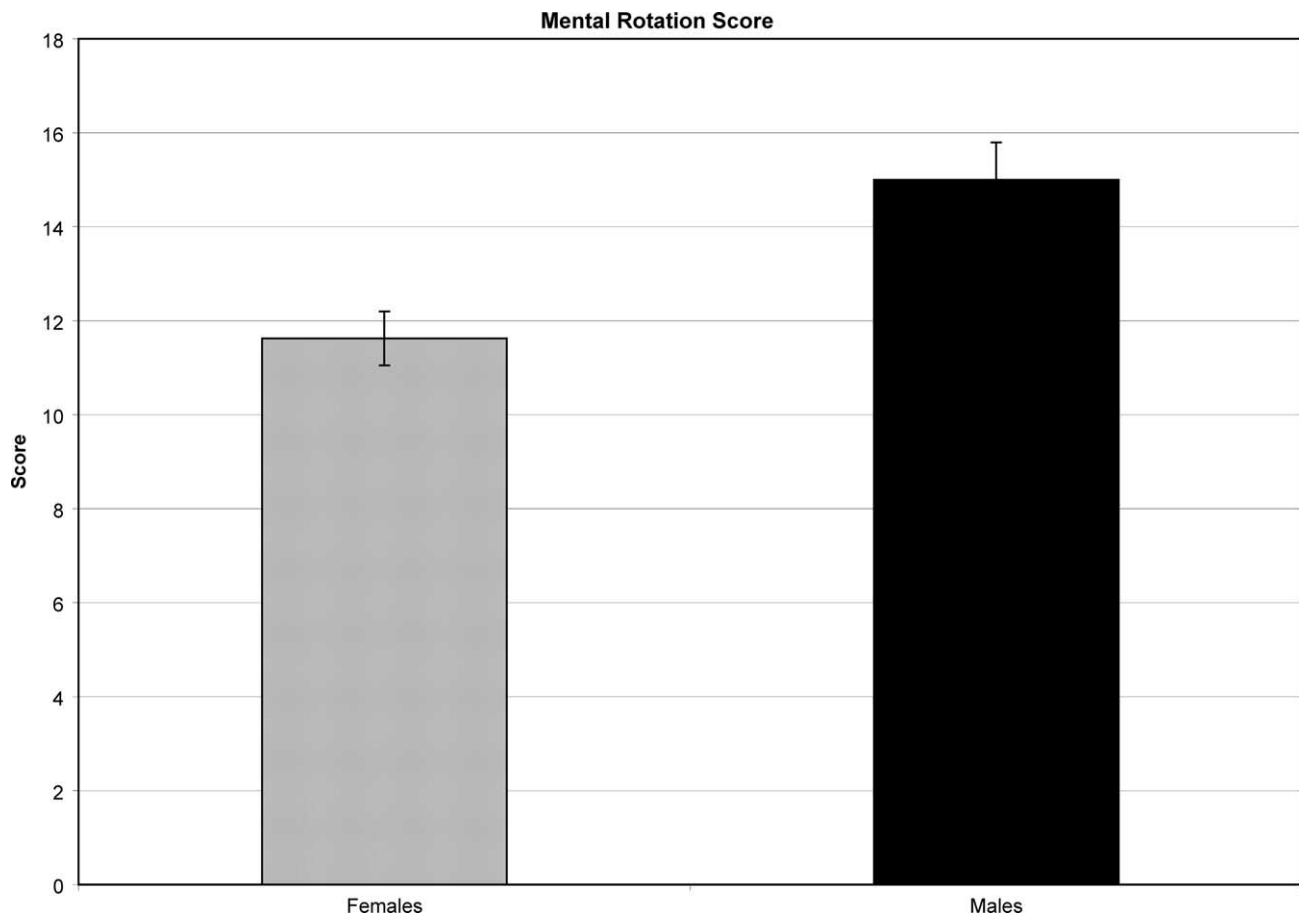


Fig. 2. Mental rotation score for males and females. Males score significantly better than females on mental rotation ( $P < 0.001$ ).

3.42,  $P < 0.001$ , with less time,  $F(19, 969) = 8.90$ ,  $P < 0.001$ , to find the platform as the number of trials increased. There were no significant interactions of Trial with Order, with Sex, or Order  $\times$  Sex. There is a significant effect of Sex on swimming latency to the hidden platform. Specifically, males swim to the hidden platform significantly more quickly than females,  $F(1, 51) = 8.74$ ,  $P < 0.01$  (Fig. 3). Examining the distance traveled to the hidden platform, there is a nonsignificant trend for males to swim a shorter distance to find the platform,  $F(1, 51) = 3.51$ ,  $P = 0.06$ . In an examination of swim speed for the training trials, there is no difference between males and females,  $F(1, 48) = 0.46$ ,  $P > 0.05$ . There was no difference between the sexes in latency to swim to the visible platform,  $F(1, 51) = 0.06$ ,  $P > 0.05$ . Examining performance on the probe trial, males spend significantly more of their distance in the training quadrant than do females,  $t(51) = 2.73$ ,  $P < 0.01$  (Fig. 4).

### 2.3. Radial arm maze

All ANOVAs are repeated measures with Trial as the repeated measure. There is a significant effect of Trial, indicating that participants traveled less distance,  $F(19, 1007) = 17.20$ ,  $P < 0.001$ , with less time,  $F(19, 1007) = 25.92$ ,  $P <$

0.001, to find the rewards as the number of trials increased. There is a significant effect of Sex on latency to find the rewards. Specifically, males find the rewards significantly more quickly than females,  $F(1, 53) = 4.83$ ,  $P < 0.05$  (Fig. 5), but there is no significant difference between the sexes in distance traveled to find the rewards,  $F(1, 53) = 1.57$ ,  $P > 0.05$  (Fig. 6). Hence, it follows that males travel significantly faster than do females,  $F(1, 53) = 15.25$ ,  $P < 0.001$ , suggesting that the sex difference in latency is simply due to males traveling more quickly than, but equally efficient to females. There is no difference between males and females in working memory errors,  $F(1, 53) = 1.05$ ,  $P > 0.05$  (Fig. 7), or in reference memory errors,  $F(1, 53) = 2.03$ ,  $P > 0.05$  (Fig. 8).

### 2.4. Order effects

Because half the participants performed the eight-arm maze first, and half performed the pool first, we can examine the effect of Order on performance. For the pool, there is a significant effect of Order in that participants are quicker to find the hidden platform if they had first had RAM training,  $F(1, 51) = 8.46$ ,  $P < 0.01$ , but there is no Sex  $\times$  Order interaction,  $F(1, 51) = 0.08$ ,  $P > 0.1$ . In addition, there is a sig-

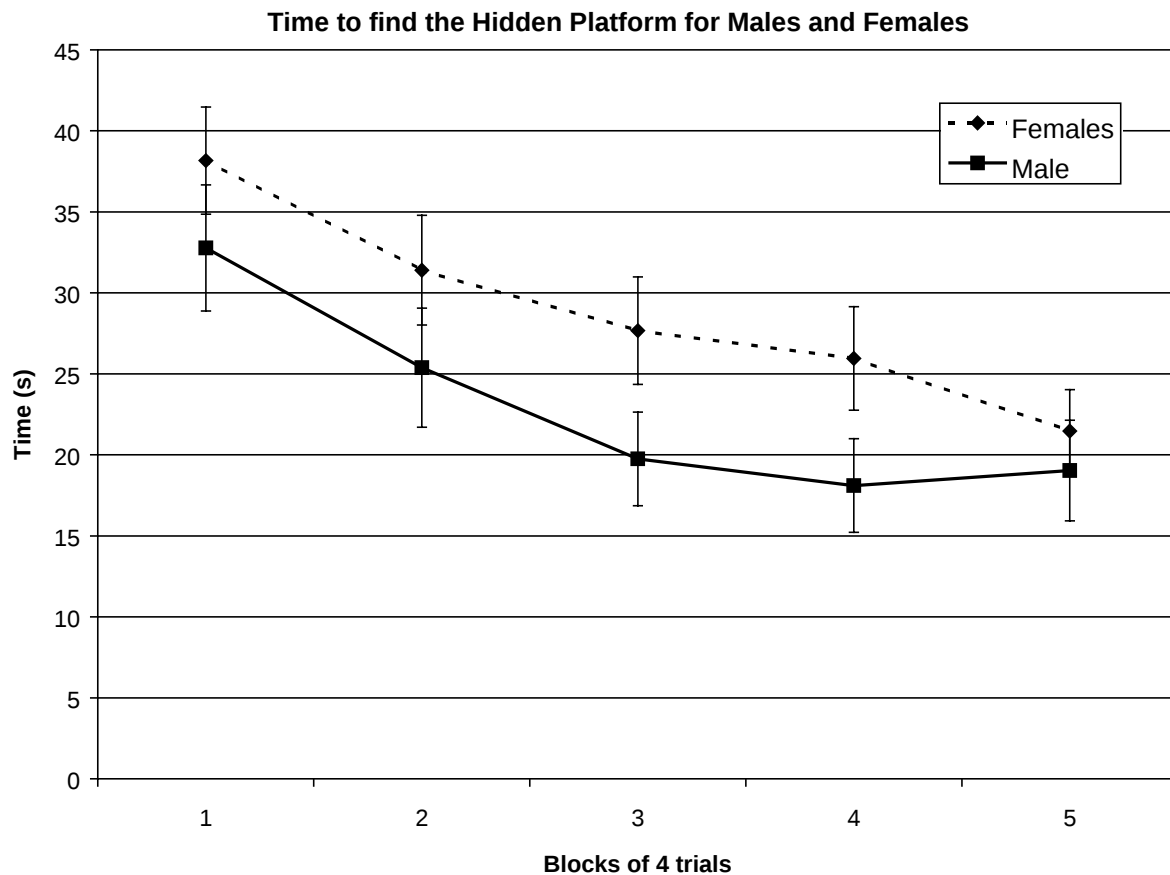


Fig. 3. Latency to find the hidden platform for males and females. Males swim to the hidden platform significantly more quickly than females ( $P < .01$ ).

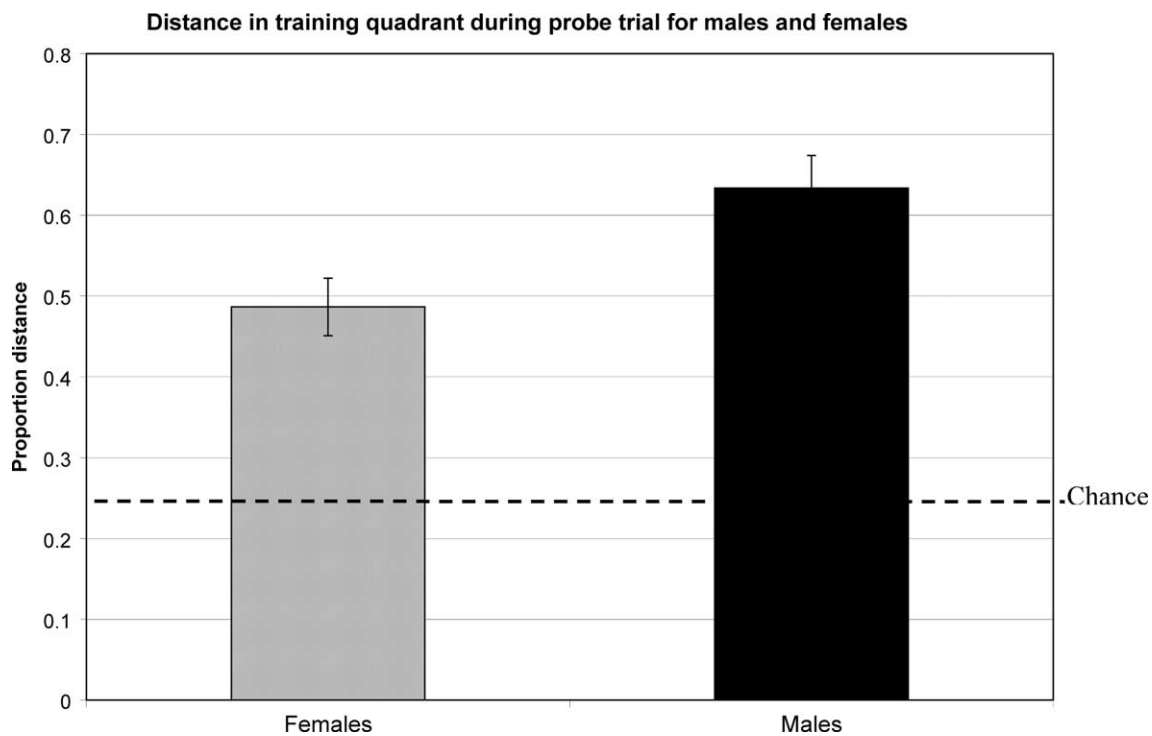


Fig. 4. Mean proportion of distance spent in the training quadrant during the probe trial. Males spend significantly more distance in the training quadrant than do females ( $P < 0.01$ ). The dashed line at 0.25 indicates chance performance.

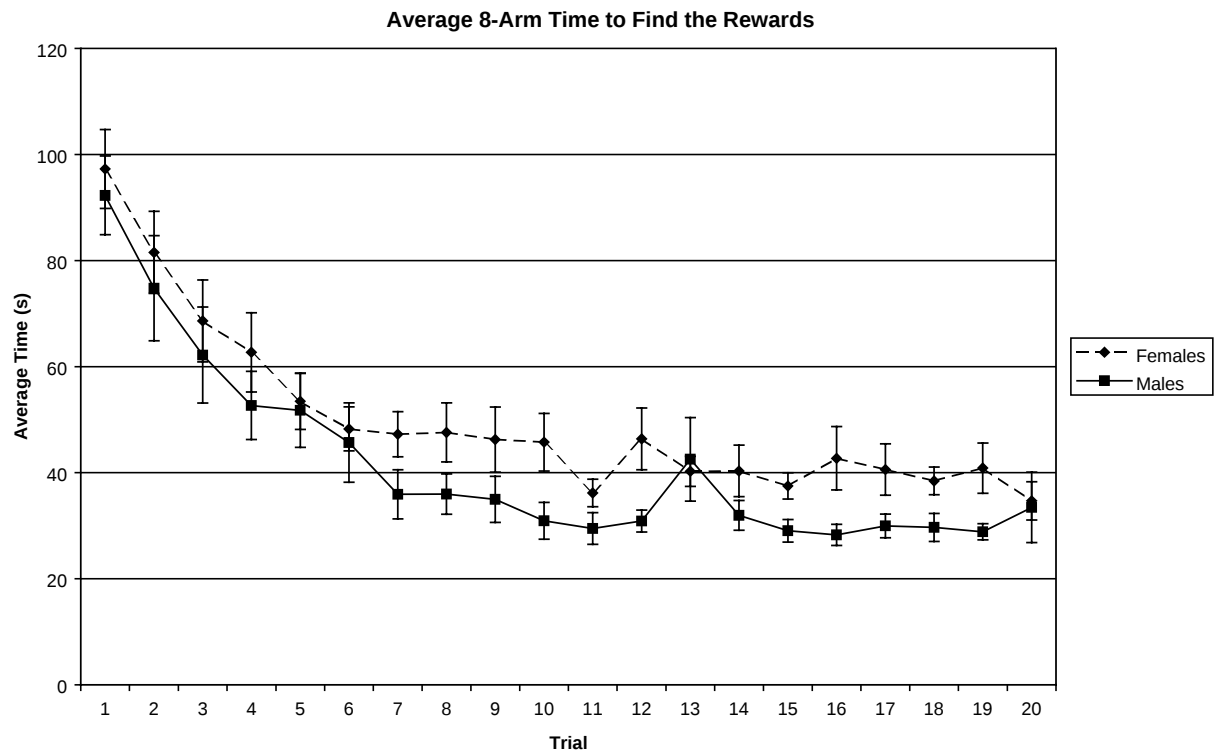


Fig. 5. Time to complete the radial arm maze for males and females. Males complete the maze significantly more quickly than females ( $P < 0.05$ ).

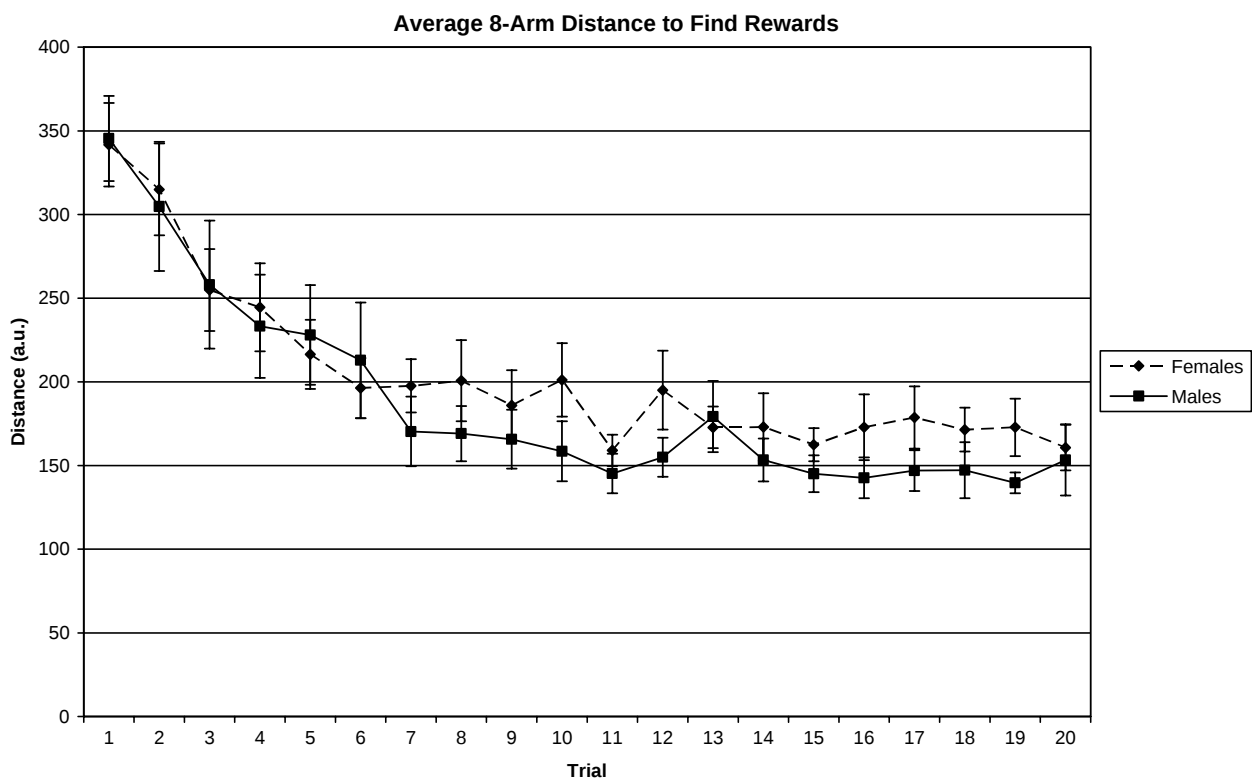


Fig. 6. Average distance to complete the radial arm maze for males and females. There is no significant difference between males and females on distance to complete the maze ( $P > 0.05$ ).

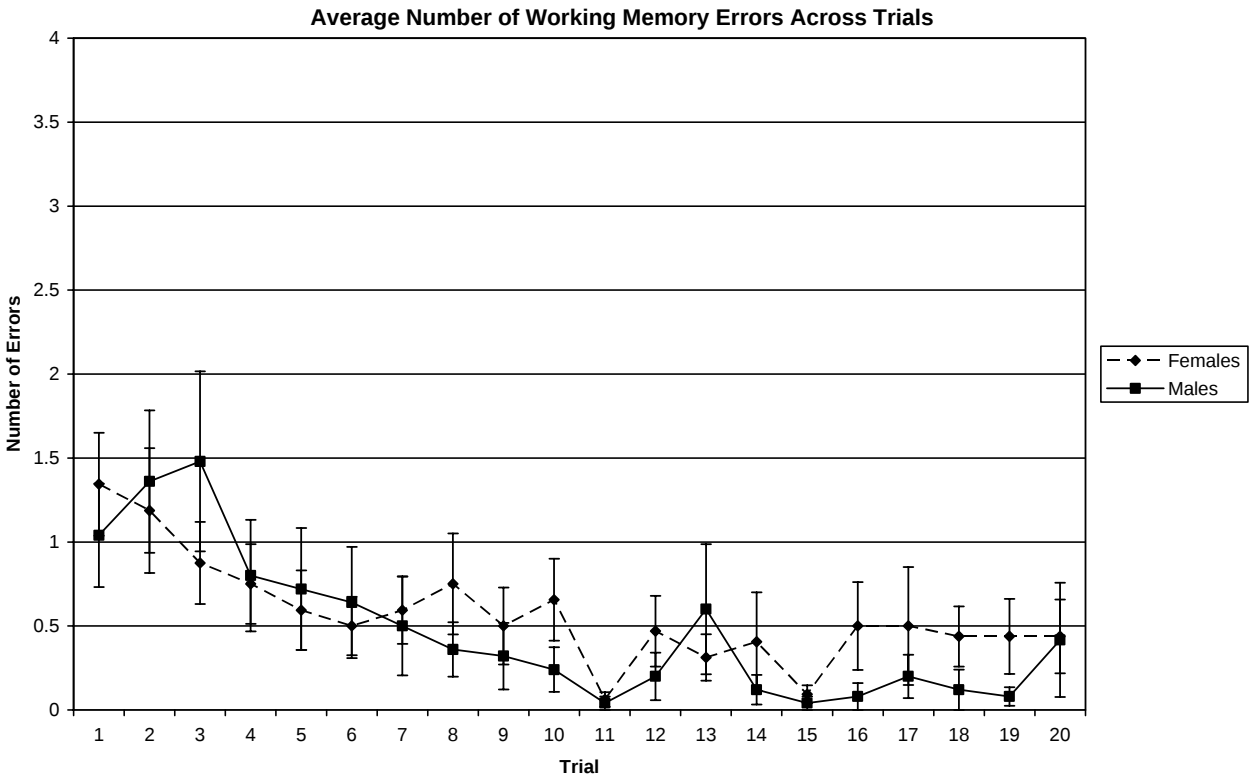


Fig. 7. Average number of working memory errors for males and females. There is no significant difference between males and females ( $P > 0.05$ ).

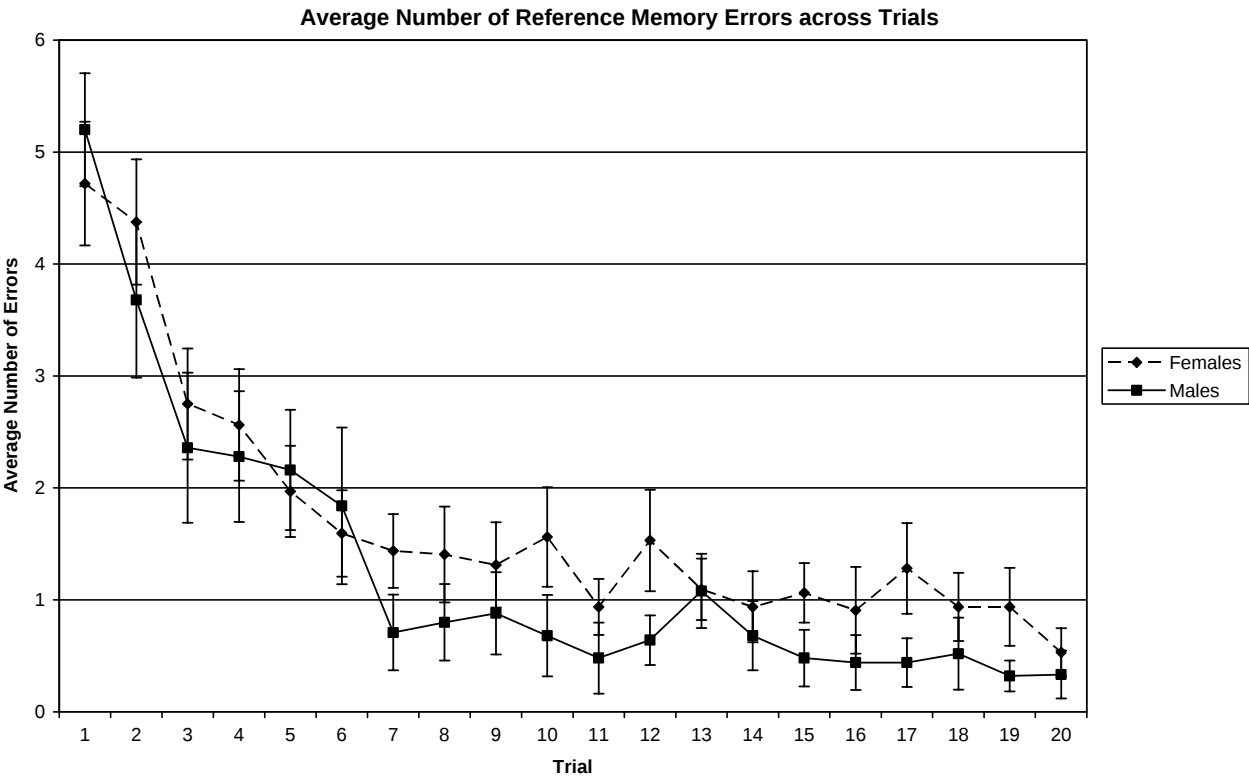


Fig. 8. Average number of reference memory errors for males and females. There is no significant difference between males and females ( $P > 0.05$ ).

nificant effect of Order in that participants travel less distance to find the hidden platform if they had first had eight-arm maze training,  $F(1, 51) = 5.42$ ,  $P < 0.05$ . For the RAM, there is a significant effect of Order in that participants are quicker to find the rewards if they had first had pool training,  $F(1, 53) = 12.69$ ,  $P < 0.01$ , but there is no Sex  $\times$  Order interaction,  $F(1, 53) = 0.09$ ,  $P > 0.05$ . There is a significant effect of Order for working memory errors,  $F(1, 53) = 5.81$ ,  $P < 0.05$ , and for reference memory errors,  $F(1, 53) = 10.96$ ,  $P < 0.01$ . Specifically, participants made more of both types of errors if they performed the RAM first. Collectively, these data suggest that participants performed better on the second task independent of what that task was.

### 2.5. Task correlations

For the correlation matrix, there are nine factors to examine: (1) average time to complete the RAM; (2) average distance to complete the RAM; (3) eight-arm working memory errors; (4) eight-arm reference memory errors; (5) mental rotation task performance; (6) average time to find the hidden platform; (7) average distance to find the hidden

platform; (8) pool probe trial percentage; and (9) 3D computer game experience. Our  $N$  fluctuated slightly because not all participants were able to complete all phases of testing. Hence, the  $N$  for the eight-arm variables is slightly higher than that for the pool.

Interestingly, the only significant correlations between performance on the eight-arm maze and on the water maze are from the pool probe trial (Fig. 9). The pool probe trial has a strong negative correlation with the time and distance to find the platform ( $r < -0.6$ ), and also has a moderately strong negative correlation with time to complete the eight-arm maze ( $r = -0.3$ ) and working memory errors on the eight-arm maze ( $r = -0.374$ ). For all of these measures, the better the performance on the probe trial (as evidenced by more amount of time swam in the training quadrant), the better the performance on the task. Additionally, we note that performance on the mental rotation task correlates negatively with both the eight-arm maze and the water maze in that those who receive high scores in mental rotation (and hence do better), also do better in the eight-arm maze and the pool as exhibited by less time and distance to complete the eight-arm maze or the pool. Moreover, those who per-

|                             |   | 8Arm<br>Time | 8Arm<br>Distance | 8Arm<br>Working<br>Memory | 8Arm<br>Reference<br>Memory | Mental<br>Rotation | Pool<br>Time | Pool<br>Distance | Pool<br>Probe |
|-----------------------------|---|--------------|------------------|---------------------------|-----------------------------|--------------------|--------------|------------------|---------------|
| 8Arm<br>Time                | r | 1            | 0.931**          | 0.824**                   | 0.930**                     | -0.348*            | 0.113        | 0.125            | -0.300*       |
|                             | p | .            | 0.00             | 0.00                      | 0.00                        | 0.01               | 0.44         | 0.39             | 0.04          |
|                             | N | .            | 52               | 52                        | 52                          | 52                 | 49           | 49               | 47            |
| 8Arm<br>Distance            | r | 0.931**      | 1                | 0.893**                   | 0.979**                     | -0.277*            | 0.054        | 0.137            | -0.2□□        |
|                             | p | 0.00         | .                | 0.00                      | 0.00                        | 0.047              | 0.715        | 0.348            | 0.071         |
|                             | N | 52           | .                | 52                        | 52                          | 52                 | 49           | 49               | 47            |
| 8Arm<br>Working<br>Memory   | r | 0.824**      | 0.893**          | 1                         | 0.799**                     | -0.319*            | 0.104        | 0.1□1            | 0.374*        |
|                             | p | 0.00         | 0.00             | .                         | 0.00                        | 0.021              | 0.478        | 0.271            | 0.010         |
|                             | N | 52           | 52               | .                         | 52                          | 52                 | 49           | 49               | 47            |
| 8Arm<br>Reference<br>Memory | r | 0.930**      | 0.979**          | 0.799**                   | 1                           | -0.243             | 0.043        | 0.13             | -0.223        |
|                             | p | 0.00         | 0.00             | 0.00                      | .                           | 0.083              | 0.77         | 0.374            | 0.131         |
|                             | N | 52           | 52               | 52                        | .                           | 52                 | 49           | 49               | 47            |
| Mental<br>Rotation          | r | 0.348*       | -0.277*          | -0.319*                   | -0.243                      | 1                  | -0.41**      | -0.3□4*          | 0.237         |
|                             | p | 0.01         | 0.047            | 0.021                     | 0.083                       | .                  | 0.003        | 0.01             | 0.109         |
|                             | N | 52           | 52               | 52                        | 52                          | .                  | 49           | 49               | 47            |
| Pool<br>Time                | r | 0.113        | 0.054            | 0.104                     | 0.043                       | -0.41**            | 1            | 0.88**           | -0.703**      |
|                             | p | 0.44         | 0.715            | 0.478                     | 0.77                        | 0.003              | .            | 0.00             | 0.000         |
|                             | N | 49           | 49               | 49                        | 49                          | 49                 | .            | 49               | 47            |
| Pool<br>Distance            | r | 0.125        | 0.137            | 0.161                     | 0.13                        | -0.364*            | 0.88**       | 1                | -0.663**      |
|                             | p | 0.39         | 0.348            | 0.271                     | 0.374                       | 0.01               | 0.00         | .                | 0.000         |
|                             | N | 49           | 49               | 49                        | 49                          | 49                 | 49           | .                | 47            |
| Pool<br>Probe               | r | -0.300       | -0.266           | 0.374*                    | -0.223                      | 0.237              | -0.703**     | -0.663**         | 1             |
|                             | p | 0.04         | 0.071            | 0.010                     | 0.131                       | 0.109              | 0.000        | 0.000            | .             |
|                             | N | 47           | 47               | 47                        | 47                          | 47                 | 47           | 47               | .             |
| 3D Game<br>Experience       | r | -0.066       | -0.103           | -0.086                    | -0.091                      | -0.099             | -0.099       | -0.057           | 0.074         |
|                             | p | 0.64         | 0.47             | 0.57                      | 0.52                        | 0.49               | 0.50         | 0.70             | 0.62          |
|                             | N | 51           | 51               | 51                        | 51                          | 51                 | 49           | 49               | 51            |

\* Correlation is significant at the 0.05 level (2-tailed)

\*\* Correlation is significant at the 0.01 level (2-tailed)

Fig. 9. Correlation chart of the three tasks.

form best on MRT also make fewer working memory errors in the eight-arm maze.

### 3. Discussion

The results from the mental rotation task show that males are better at mental rotation than females, and this is consistent with past research [7]. For the virtual Morris water task, males navigate more efficiently as evidenced by shorter times to find the hidden platform. In addition, during the probe trial, males show a stronger preference for the training quadrant than do females. However, when examining performance in the virtual RAM in the same participants, we no longer observe this behavioral difference between the sexes in either distance to find the rewards or in working or reference memory errors. Males do tend to find the eight-arm rewards more quickly than females, but this appears to be an artifact of males traveling more quickly through the RAM, and not more efficiently (recall that distance traveled is not significantly different for males and females).

A post hoc examination of the swim strategy used to search for the platform during the probe trial in the virtual Morris water task sheds light on the various strategies used by males and females. As can be seen in Fig. 10, participants could either (1) swim directly to the platform location (direct strategy); (2) swim to a certain location, orient toward a landmark, and from there, swim to the platform location (landmark strategy); (3) swim in a circle a fixed distance from the wall at the distance that the platform is located (circle strategy); or, (4) systematically swim back and forth in a crisscross manner (zigzag strategy). The direct strategy relies most heavily on spatial processing compared to the other three strategies.

An experimenter blind to the sex of the participant classified the swim strategies used during the probe trial for all participants. Strategy was determined independent of the correctness of the quadrant chosen (e.g. a subject could go

“directly” to the wrong quadrant). If it was not apparent which of the four strategies was being used, the trial was classified as unknown. As can be seen in Fig. 10, males tended to use a direct strategy, while females preferred strategies that were nonspatial or unclassifiable during the probe trial. This classification scheme serves as another measure to indicate that males are more likely to employ a spatial strategy to solve the virtual Morris water task compared to females. (There were no apparent strategy categories for the RAM, so a similar strategy analysis was not conducted on those data.) In fact, it may be that this sex difference is evident because the Morris water task is more liberal in terms of allowing different strategies to be used to solve it. For example, the circle strategy is a completely reliable method of finding the platform, and it usually results in only a few more seconds of swim time. Strategies different than a spatial strategy in the RAM, however, result in much longer latencies to complete the maze, and hence are severely discouraged. Hence, it may simply be that females prefer different strategies than males in both tasks, but that the virtual water maze permits them to maintain these strategies whereas they are indirectly forced to comply with a more spatial strategy in the RAM to avoid the longer completion times. We are currently pursuing experiments that further disambiguate the preferred strategies that are employed in these tasks.

An examination of sex differences seems to imply fundamental differences in the spatial demands of these two tasks. This is supported when the correlations among the dependent variables from these three tasks are examined. For example, the best overall predictors of virtual performance are the mental rotation task and pool probe trial performance. Specifically, those who performed well on mental rotation completed the virtual RAM more quickly, with shorter distances, and with fewer working memory errors. In addition, during the virtual water maze, they also found the hidden platform in shorter distances and with shorter latencies. Correlations with the pool probe trial are similar, except that the magnitude of correlations with the pool variables was much

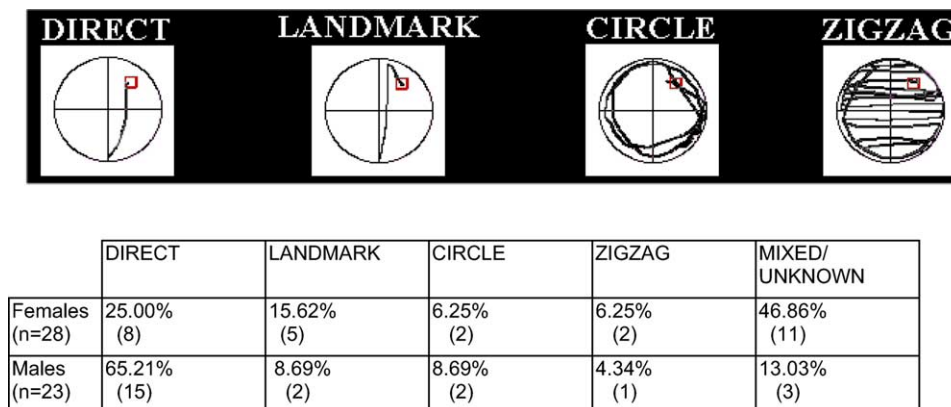


Fig. 10. Top: diagram of the four possible swim path classifications that were used to classify probe trial swims for males and females. Bottom: number of females and males that were classified in each category according to probe trial swim performance by an experimenter blind to the sex of the participants. Note that males tend to prefer a direct path during the probe while females' paths were more often a mixed use of strategies.

stronger with the probe trial, and the correlation with RAM distance did not reach quite statistical significance ( $P = 0.071$ ). However, it is interesting that there are no significant correlations between the traditional performance measures in the virtual RAM and the virtual Morris water task, such as time or distance to complete the task. This is remarkable given that both are widely used interchangeably to assess spatial memory in nonhumans. Our current data suggest that instead of these traditional measures, perhaps mental rotation performance or performance on the pool probe trial is the best single measure of spatial memory. It is not clear offhand how well these data would generalize to spatial memory performance for nonhumans for the two tasks, but probe trial performance is commonly accepted as the single best measure of spatial memory in the Morris water task. Unfortunately, there is not a “probe” test that is commonly used at the end of RAM training that may similarly encapsulate spatial memory, and of course, mental rotation ability is not something that can be easily tested in nonhumans.

It is experimentally reassuring that within each task, there are high correlations among the dependent variables such that good performance on one variable predicts good performance on the other variables. For example, in the RAM, there is a 0.93 correlation between time to complete the maze and distance to complete the maze, suggesting that they are equally valuable measures of performance in this task.

Whereas a similar lack of corroboration between the two tasks has been reported in research with rodents, these differences often were attributed to nonspatial aspects of the tasks such as the differences in stress, motivation, and motor demands of the tasks [12]. However, as mentioned previously, these procedural differences are eliminated in the virtual versions of these tasks. Specifically, whereas the actual Morris water task is an aversive motivator and is often stressful and taxing for the rodents compared to the RAM, the virtual Morris water task has similar motivation, stress, and motor requirements as the virtual RAM: hence, these factors alone cannot account for the performance differences between these two tasks.

Therefore, we should examine the procedural differences of the tasks more closely. One notable difference between the tasks is that the RAM includes a working memory component that is absent in the water maze. (Recall that in the RAM, subjects must avoid re-entering arms that they have previously visited.) However, both tasks contain a reference memory component (e.g. RAM, the location of the four baited arms; pool, the location of the platform). Since the water maze only contains a reference memory component, then it follows that a sex difference for reference memory should be observed in the RAM, if indeed reference memory is the critical factor responsible for eliciting spatial memory sex differences. However, recall that males and females do not differ on number of reference memory or working memory errors in the RAM, so this factor alone does not seem to account for the difference in performance between the tasks.

Additionally, in the RAM, it is not clear how independent the WM errors variable and RM errors variable are given that they seem to have a strong correlation with each other ( $r = 0.799$ ). In fact, working memory and reference memory may be a bit blurred in our testing protocol. With rodents, testing is performed over a number of days, so that the factors that remain constant across days are often coined “reference memory” factors while those things that may change within a single test session are coined “working memory.” However, all of our testing is performed within one session, so that working memory is not necessarily distinct from reference memory. Note that, though the idea that working memory and reference memory are distinct and independent measures is probably inaccurate even in the use of them within the rodent RAM. For example, poor spatial ability would make one unable to remember which arms they had visited in an individual session (i.e. working memory errors) as well as make one unable to locate the baited arms (i.e. reference memory errors). Hence, in the standard RAM, there may exist a number of secondary factors that correlate highly with both WM and RM. It is only through a variety of manipulations across experiments that WM and RM may be dissociated from one another (e.g. eliminating spatial behavior as a factor by using different colored or different textured arms). Either way, it should be cautioned that neither the actual RAM nor the virtual RAM can claim that WM and RM are assessed independently using the standard RAM protocols.

Another important difference between the tasks is that the RAM is more generous than the water maze in allowing error in the use of the distal cues to locate the goal locations. Specifically, in the RAM, a subject does not need to encode cue relations very precisely in order to determine which arm contains a goal. In fact, it is quite possible to encode only one cue at the distal end of an arm and not to attend to cue relations at all. For the eight-arm RAM, each arm subtends a  $45^\circ$  angle, and hence there is little distal cue overlap within one’s field of view that makes it necessary to attend to relations among multiple cues. Additionally, there are only eight possible routes to take to the goal in the RAM, whereas the number of possible routes in the water maze is much larger since there are no confined passageways.

Note additionally that the both with the real versions and virtual versions, the RAM simply requires knowledge of a direction within the maze, whereas the water maze requires knowledge of both direction and distance. For example, once an animal enters an arm of the RAM, the length of that arm is irrelevant: the animal knows it must travel to the distal end of the arm. Additionally, once this arm is entered, the goal at the end of the arm is visible and no further localization of the goal is required. In contrast, in the water maze, the goal is always hidden, and direction is accompanied by a judgment of distance to locate the platform. Theoretically, an accurate directional knowledge should be sufficient to lead the animal to encounter the platform, but as has been shown by Whishaw et al. [38], rodents typically slow down

as they approach the platform, hence indicating knowledge of its precise distance.

Another difference between the tasks is that whereas there are typically four different start locations in the water maze, there is only one start location in the RAM, but the animal is pointed in different start directions. This may allow the animal to orient itself better within the RAM environment and hence may result in a more consistent search pattern than in the water maze.

These task differences notwithstanding, task difficulty alone cannot explain the performance discrepancies in these two tasks. Recall that there are reports of spatial navigation impairments in the MWM and not the RAM following ischemia in rats [22], but the opposite pattern of results is observed in other studies [10]. Hence, there seem to be other factors contributing to these differences besides task difficulty.

In deciphering the sex difference in the virtual Morris water task, we note that Sandstrom et al. [31] report that males are particularly attuned to the room geometry whereas females are more likely to attend to landmark information during virtual navigation. This may suggest in our experiment that since two different virtual rooms were used for the two mazes, it may be that the water maze was simply situated in a room more likely to reveal a sex difference (e.g. in a room that had different geometrical information which the males could use to their advantage). However, the RAM used in this experiment is in the exact same virtual room and in the same location as the virtual pool used previously in experiments that reveal very robust sex differences in a virtual Morris water task [2,15]. Hence, it does not seem that room geometry alone can account for this difference between the tasks.

Note that our results do not differentiate between strategy ability and strategy preference; it may be that females can navigate as efficiently as males, but that they prefer to use different strategies. Moreover, there are a variety of other factors that may contribute to these performance differences. For example, Lawton [17] has reported a negative correlation with spatial performance and spatial anxiety in similar paradigms, and Quaiser-Pohl and Lehmann (2002) [29] have shown that sex differences in mental rotation are of a greater magnitude in students studying arts, humanities, and social sciences, and smallest for those students majoring in computational visualistics. Furthermore, Dabbs et al. [9] have also show that independent of sex, older subjects gave more abstract Euclidian directions than did younger subjects. Additionally, it has been shown that females perform better on the mental rotation task during menstruation (when estrogen was lowest) as opposed to during their luteal phase (when estrogen is highest) [19,32]. Hence, each of these factors, alone or in combination, probably contribute to strategy preferences and abilities in navigation and mental rotation, and future experimentation is necessary to disambiguate the role of these factors in sex differences and performance.

Lastly, it is not clear whether the use of virtual environments is an accurate analogue of actual navigation. For ex-

ample, in actual navigation, the participant receives various sensory and vestibular feedback which are largely absent in virtual environments. However, it should be noted that vestibular stimulation derived from optic flow still is present in virtual environments, as evidenced by a small portion of participants reporting motion sickness during the tasks. More importantly, these procedural differences aside, males and females do not differ on swimming to a visible platform, or in the RAM, so factors such as vestibular feedback, joystick manipulation, and interactions with the 3D program are similar between the sexes and hence cannot account for the observed differences.

Collectively, these data are the first to utilize a virtual eight-arm task, and also the first to compare these two tasks in humans. Interestingly, we observe that males perform better than females in the virtual water maze, but that there is no sex difference in spatial performance in the same participants when performing a virtual RAM. Moreover, the correlations between the two virtual tasks indicate that good performance on one virtual task does not significantly correlate with good performance on the other virtual task. We believe that these differences are due to strategic or working memory demands that differ between the tasks, and that various factors such as motivation, movement, and stress cannot fully account for these differences since these factors were equivalent in both tasks. Hence, our data show that, as in rodents, these two tasks reveal different spatial memory sensitivities of the water maze and the RAM, and caution should be taken when employing these tasks interchangeably as tests of spatial memory.

## Acknowledgements

Funded by NSF IBN 9809958, NIH NS-40497, and NIH 1 F32 MN64290-01.

## References

- [1] Aguirre GK, Detre JA, Alsop DC, D'Esposito M. The parahippocampus subserves topographical learning in man. *Cereb Cortex* 1996;6:823–9.
- [2] Astur RS, Ortiz ML, Sutherland RJ. A characterization of performance by men and women in a virtual Morris water task: a large and reliable sex difference. *Behav Brain Res* 1998;93:185–90.
- [3] Astur RS, Constable RT. Hippocampal activation during a virtual Morris water task and transverse patterning. *Soc Neurosci Abstr* 2001;74:12.
- [4] Astur RS, Taylor LB, Mamelak AN, Philpott L, Sutherland RJ. Humans with hippocampus damage display severe spatial memory impairments in a virtual Morris water task. *Behav Brain Res* 2002.
- [5] Bingman V. Unimpaired acquisition of spatial reference memory, but impaired homing performance in hippocampal ablated pigeons. *Behav Brain Res* 1988;27:179–88.
- [6] Bohbot VD, Kalina M, Stepankova K, Spackova N, Petrides M, Nadel L. Spatial memory deficits in patients with lesions to the right hippocampus and to the right parahippocampal cortex. *Neuropsychologia* 1998;36(11):1217–38.

- [7] Collins DW, Kimura D. A large sex difference on a two-dimensional mental rotation task. *Behav Neurosci* 1997;111(4):845–9.
- [8] Corkin S. Tactually-guided maze learning in man: effects of unilateral cortical excisions and bilateral hippocampal lesions. *Neuropsychologia* 1965;3:339–51.
- [9] Dabbs JM, Chang EL, Strong RA, Milun R. Spatial ability, navigation strategy, and geographic knowledge among men and women. *Evol Human Behav* 1998;19:89–98.
- [10] Davis HP, Volpe BT. Memory performance after ischemic or neurotoxin damage of the hippocampus. In: Squire LR, Lindenlaub E, editors. *The biology of memory*. In: *Symposia Medica Hoechst* 23, FK. Stuttgart: Schattauer Verlag; 1990. p. 477–507.
- [11] Gron G, Wunderlich AP, Spitze M, Tomczak R, Riepe MW. Brain activation during human navigation: gender-different neural networks as substrate of performance. *Nat Neurosci* 2000;3(4):404–8.
- [12] Hodges H. Maze procedures: the radial-arm and water maze compared. *Cog Brain Res* 1996;3:167–81.
- [13] Jacobs WJ, Laurance HE, Thomas KGF. Place learning in virtual space I: acquisition, overshadowing, and transfer. *Learn Motiv* 1997;28:521–41.
- [14] Jacobs LF, Gaulin SJC, Sherry DF, Hoffman GE. Evolution of spatial cognition: sex-specific patterns of spatial behavior predict hippocampal size. *Proc Natl Acad Sci USA* 1990;87:6349–52.
- [15] Koerner A, Astur RS, Parker E, Elwell A, Hodge GK. Menstrual cycle influences on virtual navigation and object location. *Soc. Neurosci. Abstr.* 74.5.
- [16] Krebs JR, Sherry DF, Healy SD, Perry VH, Vaccarino AL. Hippocampal specialization of food-storing birds. *Proc Natl Acad Sci* 1989;86(4):1388–92.
- [17] Lawton CA. Gender differences in way-finding strategies: relationship to spatial ability and spatial anxiety. *Sex Roles* 1994;30:765–79.
- [18] Maguire EA, Burgess N, Donnett JG, Frackowiak RSJ, Frith CD, O'Keefe J. Knowing where and getting there: a human navigation network. *Science* 1998;280:921–4.
- [19] Moody SM. Changes in scores on the mental rotations test during the menstrual cycle. *Percept Motor Skills* 1997;84(1):955–61.
- [20] Morris RGM. Spatial localization does not require the presence of local cues. *Learn Motiv* 1981;12:239–69.
- [21] Morris RGM, Garrud P, Rawlins JP, O'Keefe J. Place navigation in rats with hippocampal lesions. *Nature* 1982;297:681–3.
- [22] Nunn JA, Hodges H. Cognitive deficits induced by global cerebral ischaemia: relationship to brain damage and reversal by transplants. *Behav Brain Res* 1994;65:1–31.
- [23] O'Keefe JA, Dostrovsky J. The hippocampus as a spatial map: preliminary evidence from unit activity in freely moving rats. *Brain Res* 1971;34:171–5.
- [24] O'Keefe JA, Nadel L. *The Hippocampus as a cognitive map*. Oxford: Clarendon Press; 1978.
- [25] Oler JA, Markus EJ. Age-related deficits on the radial maze and in fear conditioning: hippocampal processing and consolidation. *Hippocampus* 1998;8:402–15.
- [26] Olton DS, Samuelson RJ. Remembrance of places passed: spatial memory in rats. *J Exp Psychol: Anim Behav Proc* 1976;2:97–116.
- [27] Olton DS, Walker JA, Gage FH. Hippocampal connections and spatial discrimination. *Brain Res* 1978;139:295–308.
- [28] Perrot-Sinal TS, Kostenuik MA, Ossenkopp KP, Kavaliers M. Sex differences in performance in the Morris water maze and the effects of initial nonstationary hidden platform training. *Behav Neurosci* 1996;110:1309–20.
- [29] Quaiser-Pohl C, Lehmann W. Girls' spatial abilities: charting the contributions of experiences and attitudes in different academic groups. *Brit J Educ Psychol* 2002;72:245–60.
- [30] Roof RL. Neonatal exogenous testosterone modifies sex difference in radial arm and Morris water maze performance in prepubescent and adult rats. *Behav Brain Res* 1993;53:1–10.
- [31] Sandstrom NL, Kaufman J, Huettel SA. Males and females use different distal cues in a virtual environment navigation task. *Cog Brain Res* 1998;6(4):351–60.
- [32] Silverman I, Phillips K. Effects of estrogen changes during the menstrual cycle on spatial performance. *Ethol Sociobiol* 1994;14(4):257–69.
- [33] Smith ML, Milner B. The role of the right hippocampus in the recall of spatial memory. *Neuropsychologia* 1981;19:781–93.
- [34] Sutherland RJ, Whishaw IQ, Kolb B. Spatial mapping: definitive disruption by hippocampal and frontal cortex damage in the rat. *Neurosci Lett* 1982;31:271–6.
- [35] Tolman EC. Cognitive maps in rats and men. *Psychol Rev* 1948;55: 189–208.
- [36] Vanderberg S, Kuse AR. Mental rotations: a group test of three-dimensional spatial visualization. *Percept Motor Skills* 1978;47:599–604.
- [37] Wertlieb D, Rose D. Maturation of maze behavior in preschool children. *Develop Psychol* 1979;15(4):479–80.
- [38] Whishaw IQ, Cassel JC, Jarrard LE. Rats with fimbria-fornix lesions display a place response in a swimming pool: a dissociation between getting there and knowing where. *J Neurosci* 1995;15:5779–88.