



A Field-Based Adaptation of the Classic Morris Water Maze to Assess Learning and Memory in a Free-Living Animal

Leah McKay¹, Louis Hunninck^{1,2}, and Michael J Sheriff^{1,*}

¹ Biology Department, University of Massachusetts Dartmouth

² Department of Natural Resources and Environmental Sciences, University of Illinois at Champaign-Urbana

*Corresponding author (Email: msheriff@umassd.edu)

Citation – McKay, L., Hunninck, L., & Sheriff, M. J. (2022). A field-based adaptation of the classic Morris Water Maze to assess learning and memory in a free-living animal. *Animal Behavior and Cognition*, 9(4), 396-407. <https://doi.org/10.26451/abc.09.04.03.2022>

Abstract – Free-living animals rely on cognitive functions, such as learning and memory, for numerous aspects of their survival. However, research involving these mechanisms is often limited to the laboratory where animals are far removed from natural influences. Here, we adapted the Morris Water Maze (MWM) for use in the field to test learning and memory of free-living white-footed mice, *Peromyscus leucopus*. The MWM consists of a pool filled with opaque water and a platform that is hidden beneath the surface of the water. Mice were tasked with finding an escape platform using proximate-cue based spatial learning skills across trials within and between two tests. The two tests occurred 2 hrs apart and each consisted of 5 trials with a maximum search time of 60 s, separated by 15-s breaks. We found mice demonstrated rapid learning within the first 5 trials, as indicated by a faster time and reduced distance swam to find the platform in trial 5 as compared to trial 1 in Test 1. We also found mice remembered the task, as indicated by equivalent performance (in time and distance) after the two-hour delay period (i.e., performance in Test 2 trial 1 was similar to that in Test 1 trial 5). Lastly, we found that there was no increase in performance as mice completed an additional 5 trials (i.e., no difference in Test 1 trial 5 vs. Test 2 trial 5). By assessing several variables that serve as standard metrics for learning and memory within the MWM, we were able to illustrate the ability of wild mice to successfully navigate within a modified version of an established learning/memory paradigm. This study is important for promoting more accessible forms of cognitive testing in free-living animals and use of such a method could provide insights into how animals cope with environmental stressors in their natural environment such as anthropogenic disturbance, invasive species, and habitat degradation. A better understanding of cognitive behavior in wild animals may help diminish the disparity of knowledge that is often present between laboratory and field research.

Keywords – Cognition, Cognitive ability, Cognitive behavior, *Peromyscus leucopus*, White-footed mouse

Cognitive functions, such as learning and memory, are essential for animals' ability to forage, reproduce, defend a territory, and avoid predators (Mastrangelo et al., 2009). Learning is defined as a neural process that uses previous experiences to adjust behavior, whereas, memory is the process of storing learned information so that it can be accessed later (Brocardo & Gil-Mohapel, 2012). Learning is a general term that can be divided into more specific forms including spatial learning (the acquisition of spatial information), associative learning (the links between cues and rewards), or motor learning (acquisition of information pertaining to action sequences; Morand-Ferron et al., 2016). Memory systems can be classified into two broad categories: declarative (explicit) and non-declarative (implicit); declarative memory involves a conscious recollection of information, while non-declarative is more of an umbrella term that encompasses forms of memory that are accessed without conscious recollection (Squire, 1992). Memory can also be classified into long-term, short-term, and working memory. Long-term memory is a vast storage of knowledge on previous events. Short-term memory holds a limited amount of information accessible in

a temporary state. Finally, working memory is similar to short-term memory in that it is the temporary storage of information, however, this information is retained and used to plan and carry out behaviors (Brocardo & Gil-Mohapel, 2012).

The ability to learn and memorize has been conserved throughout the evolution of animals and has arisen in response to the challenges that they face in their natural environment (Morand-Ferron, 2017). As such, cognitive behavior is assumed to be essential for animals to cope with the environment they encounter (Cauchoix et al., 2017; Mastrangelo et al., 2009). Yet, nearly all the studies examining cognitive behavior (i.e., learning and memory) come from laboratory or captive-based studies. Spatial and associative learning and memory may be of particular importance in allowing wild animals to cope with and respond to stressors within their environment, however, interestingly, many laboratory-based studies find that such stressors reduce individual cognitive behavior (e.g., Adamec et al., 2004, 2006; Diamond et al., 2006; Park et al., 2008). Thus, there is a need to better understand cognitive behavior in free-living wild animals. In addition, testing cognitive behavior in free-living animals will reduce the need for housing or quarantine considerations of wild animals, reducing the disruption of the individual from its natural environment and/or allowing testing of animals that may not be able to be maintained in laboratory or captive settings. For example, if a study was interested in comparing urban vs. rural environments, individuals could remain within the environment they inhabit throughout testing.

While there are numerous tests for assessing spatial learning and memory in animals, including the radial arm maze (Olton & Samuelson, 1976), the Barnes maze (Barnes, 1979), and the Elevated Plus maze (Handley & Mithani, 1984), one of the most widely employed methods is the Morris Water Maze, originally described by Richard Morris (Morris, 1984). The Morris Water Maze (MWM) was initially designed to investigate spatial learning and memory in laboratory rats. In its most basic form, this apparatus consists of a large pool filled with opaque water and a platform that is hidden beneath the surface of the water. A camera is suspended above the pool to record movement within the maze. Subjects placed within the pool are tasked with finding an escape platform using spatial learning skills. This original design does not include any proximal cues, so rats within the maze rely on distal cues to recognize the location of the platform. Laboratory animals usually become sufficiently trained in the task within 5-6 days with numerous trials on each day or in 10-12 days with a single trial on each day (Vorhees & Williams, 2006). Advantages to using the MWM include its simplicity and ease of testing (Vorhees & Williams, 2014), but also its ability to be modified to assess various other types of learning and memory. This includes testing reversal learning where the platform is relocated to the opposite side of the pool after training, discrimination learning using multiple platforms, and nonspatial learning, cue learning, and sensorimotor function through the use of additional cues (Morris, 1984; Vorhees & Williams, 2006).

One learning test the MWM is commonly modified for is that of proximal associative-cue spatial learning and memory (Anisman & McIntyre, 2002; Baldan Ramsey & Pittenger, 2010; Upchurch & Wehner, 1988; Vorhees & Williams, 2006). In this modified version of the MWM animals learn to associate a proximate cue (a beacon; Shettleworth, 1998) with the escape platform; they learn similar skills as they do during the training period of the distal-cued spatial version of the MWM (Fukuda et al. 2002; Hoh et al., 1999; Vorhees & Williams, 2014). While distal cues are often the focus of research involving spatial learning, research suggests that the use of proximal, beacon cues is similarly important for spatial navigation (Carman, 2002; Hébert et al., 2017; Shettleworth 1998; Vollmer-Conna & Lemon, 1998) and can provide a reliable metric for assessing search strategies, associative/ conceptual abilities, and working memory (Anisman & McIntyre, 2002); however, it must be noted that the importance of local (beacon) vs global (landmark) cues can vary by species, size of stimuli, and individual experience (Bennett, 1993; Cheng, 1995; Dyer, 1987; Vlasak, 2006). One of the major advantages to the proximal cue version of the MWM is that it does not require an extensive training period. While this may not be a problem in the laboratory (where MWM test are routinely and have thus far been exclusively used), it has inhibited the use of this technique in field studies, where holding wild animals over multiple days is rarely feasible and would disrupt their normal activity, thus, eliminating the intent of examining cognitive ability in free-living animals.

There is significant value in understanding the cognitive abilities of free-living animals that have evolved multiple responses to cope with the ecological challenges found within their natural environments (Cauchoix et al., 2017). Yet few studies have explored such processes, possibly due to the limitations involved in testing (e.g., the training period, Morand-Ferron et al., 2016). Here, we adapted the proximal cue version of the MWM (Vorhees & Williams, 2006) for use in the field with free-living white-footed mice, *Peromyscus leucopus*. This version eliminates the need for multiple day training that would require wild animals be held in captivity or re-trapped across multiple days. Our modified version consists of two test periods, each with 5 consecutive trials (10 in total) and a 2 hr rest period between tests, in a single day. We predicted that 1) mice would perform better after 5 trials compared to the naïve trial (indicating learning), 2) mice would retain that better performance after a 2 hr rest period (indicating memory), and 3) mice would have similar performance after 10 trials as they had after the first 5 trials (indicating that mice had reached their best potential performance, and that our protocol measures learning sufficiently). Mice performance during trials was evaluated using 3 different metrics: time to find platform, distance traveled to find the platform, and time spent in the target quadrant (i.e., quadrant of pool containing platform). Better performance is indicated by 1) reduced time to find the platform, 2) a shorter distance to find the platform, and 3) more time in the target quadrant of the pool.

Method

Ethics Statement

Animals involved were cared for in accordance with the University of Massachusetts Dartmouth IACUC guidelines, who reviewed and approved methods used (IACUC #19-2). Individuals were provided with food and bedding while in the trap and released at the original capture location in good condition.

Study Area

This study was conducted in the Southeastern Massachusetts Bioreserve (41°44'21.9"N 71°01'11.2"W) on state managed land. The area is made up of upland forests of mixed hardwoods and pine with shrub understory and is approximately 2225 hectares. Dominant tree species include eastern white pine (*Pinus strobus*), mixed oaks (*Quercus spp.*), pitch pine (*Pinus rigida*), red maple (*Acer rubrum*), and black birch (*Betula lenta*). Understory vegetation is mainly low bush blueberry (*Vaccinium angustifolium*), bracken fern (*Pteridium aquilinum*), black huckleberry (*Gaylussacia baccata*), and greenbrier (*Smilax spp.*). Research was conducted on six 150 m X 150 m trapping grids (Figure 2.1) with stations 15 m apart and traps set at every other station, totaling 50 traps per grid. Each grid was located a minimum of 500 m apart to eliminate movement of mice between grids. Grids were a minimum of 200 m from any major outline through the forest.

Trapping

Live-trapping of white-footed mice took place during the summer months of 2019 and 2020. Standard trapping techniques were employed (Krebs et al., 2011) using Longworth (Penlon Ltd.) small mammal traps. Traps were pre-baited with millet for at least two days prior to trapping. At dusk, traps were set with millet, carrot, and cotton for bedding then checked at dawn, 8-10 hr later. At each capture, mice were weighed, their sex and reproductive condition were determined, and they were tagged with a unique numeric identifier. Mice were moved to a common testing area approximately 100 m from the grid (different testing areas were set up for each grid). Trapping did not occur on days with rain.

Cognitive Testing

We adapted a proximal-cue version of the Morris Water Maze to test cognitive ability in the field with free-living mice ($n = 77$ out of 90 individual mice successfully completed trials). This set-up is extensively used in laboratory studies of learning and memory (Alcalá et al., 2020; Barnhart et al., 2015; Barry & Commins, 2019; Morris, 1984; Vorhees & Williams, 2006). The maze consisted of a small pool (Diameter = 1.5 m; Circumference \approx 4.7 m; 0.5 m high walls) filled with approximately 15 cm of water turned opaque with powdered milk. A small, 60 cm² platform was hidden just below the water level so that the mice were able to climb onto it to avoid the water but not able to see it while swimming. A 15 cm high wire with a small flag was attached to the platform to act as a cue. Mice participated in two tests with five trials each. Each trial lasted a maximum of 60 s where mice needed to learn to associate the flag with the platform location. If the platform was located by the mouse, it was allowed to rest there for 15 s to provide positive reinforcement through a release from swimming. Mice that did not succeed in locating the platform within 60 s were gently placed on the platform to rest for 15 s (Vorhees & Williams, 2006). Between each trial, mice were removed from the arena for 15 s. The platform was moved among trials, as well as the location of entry for the mouse. The mice were left to rest in their holding containers with fresh cotton, apple, and millet for two hours between tests and for two hours after the final test prior to release. A video camera (GoPro Hero7) was suspended above the pool to record each trial and the researchers were out of view of the mouse during trials. We measured time to find platform, distance traveled, and proportion of time spent in the target quadrant (quadrant where platform was located). All tests were only performed on days without rain and when ambient daytime temperatures were above 22°C.

Video Analysis

To track mouse movement in the recorded videos, a computer vision approach was used. We developed a specific tracking software in Python (v. 3.7.4) using the user interface PyCharm (v. 2020.3). The main Python library used for this software was Open CV (v. 4.4.0.44, 2015, Open Source Computer Vision Library.). Other libraries used include pandas (v. 1.1.3) and numpy (v. 1.19.2). The software is based on a tracking algorithm known as Discriminative Correlation Filter with Channel and Spatial Reliability or CRST abbreviated. The CRST tracking algorithm was developed by LuNežič et al. (2018). Full code is available in supplementary info; the program packaged for Mac OS is available on GitHub. This tracking software required the researcher to manually determine the location of the pool wall (by clicking on 3 different locations on the wall), the flag position, and the starting location of the mouse before the tracking was initiated. After the tracking program was started, the x-y location of the mouse was determined for every frame of the video (framerate = 24 fps).

Time to platform was determined from the start of the video until the mouse was on the platform. Distance travelled was calculated as the total distance swum during the experiment. To reduce error in distance estimates, distances were calculated every 0.5 s, i.e., the distance between the location of the mouse in frame n and the location of the mouse on frame $n+12$. The sum of these distances then resulted in a total distance travelled. Time spent in target quadrant was determined as the proportion of total location fixes that were located within the quadrant of the pool where the flag and platform were located.

Statistical Analysis

To assess learning, we compared mice performance after 5 trials (Test 1 trial 5 (T1t5)) compared to the naïve trial (Test 1 trial 1 (T1t1)). To assess memory, we compared mice performance between T1t5 and their performance after a two-hour rest period (T2t1). To assess whether our protocol performed adequately (i.e., captured mice maximal performance), we compared mice performance in T1t5 to the performance after 5 more consecutive trials (T2t5). Lastly, to assess whether mice still improved their performance after the 5 initial trials, (i.e., Test 1) we compared mice performance in T2t1 to the performance in T2t5. These comparisons were performed for each performance metric.

We applied linear mixed effects regression analyses to each of the three performance metrics (response variable) using data of only successful trials; unsuccessful trials, where a mouse did not reach the platform within the allotted time of 60 s were discarded as mice were picked up and gently placed on the platform and the experiment was therefore interrupted (i.e., their actual time/distance to platform was not known, Vorhees & Williams, 2006). These continuous variables were square root transformed in the models to obtain normal distribution of model residuals. Additionally, residuals of all models were checked for heteroskedasticity or other patterns; no patterns were found. The fixed effect in each of these three models was a categorical variable with four levels: 1) T1t1, 2) T1t5, 3) T2t1, and 4) T2t5, and with mouse ID (i.e., unique identifier for each captured mouse) and date (to correct for differences between experiment days) as the random effect.

All analyses were conducted with the *lmer* function (R package *lme4*, v.1.1-27.1) using the *emmeans* function (R package *emmeans*, v.1.7.0) for posthoc analyses, in R (v.4.1.1) through RStudio (v.1.4.1717). Pairwise comparisons used the Tukey method for p-value adjustment due to multiple comparisons.

Results

Filtering Recordings

338 video files were considered for analyses, of which 20 were excluded due to failed trials (e.g., mouse displayed drowning behavior, mice were too difficult to track by the program, or video recordings were incomplete). Another 10 video recordings were removed after tracking due to failure of the tracking program to produce useful output (e.g., mice initially tracked well but were subsequently lost by the tracking program). On a closer look of the 308 video recordings that passed the previous filters, some recording did not completely record the mouse movements but stopped recording during the experiments ($N = 65$). We were thus left with 243 (72%) video recordings for analysis.

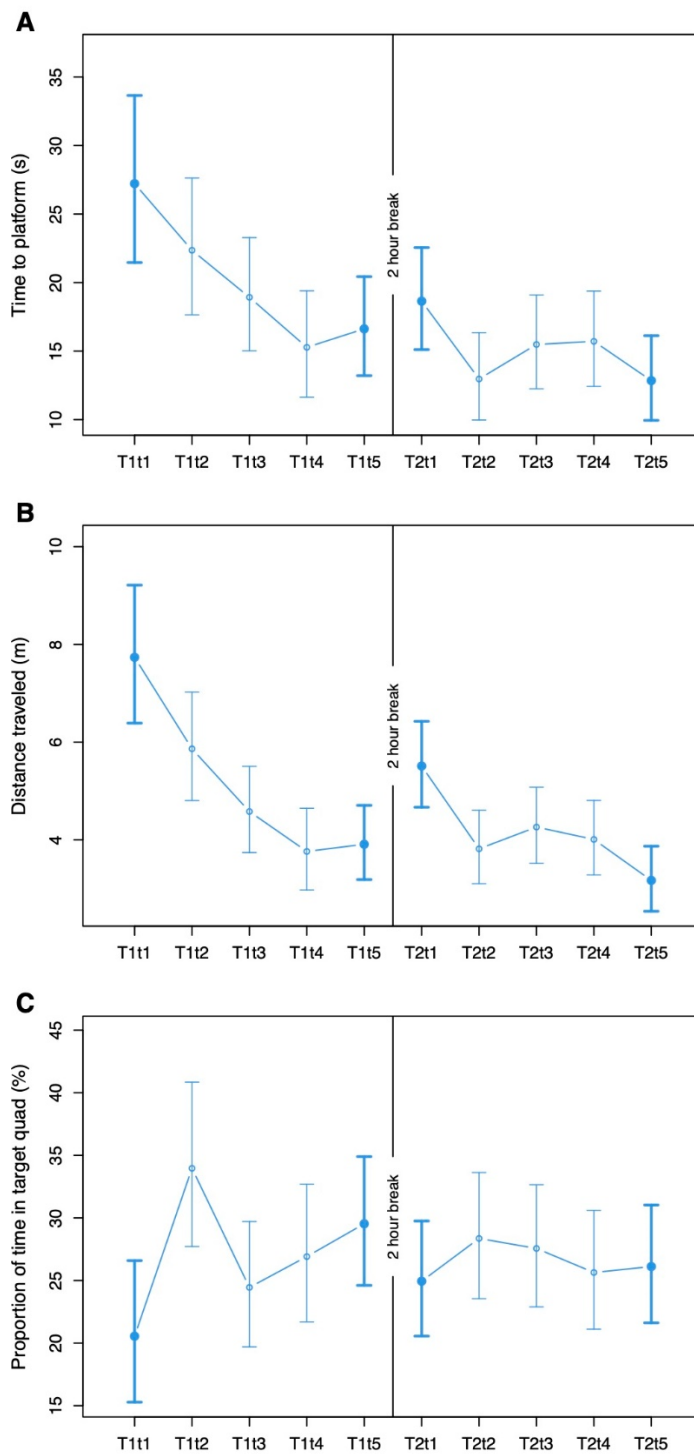
Mice in the Experiment

Of the 243 trials, mice in 44 (18%) did not reach the platform within the predetermined end of one experiment of 60 s. Given the mice were placed on the platform at this time and the trial stopped, they were excluded from the analyses, which left us with 199 successful trials. Thus, given the removal of these specific trials, mice safety (i.e., drowning behavior), and the tracking difficulties certain trials (and mice) were removed from our analysis; overall there were 31 successful trials of a total of 46 in T1t1, 54 of 67 in T1t5, 57 of 59 in T2t1, and 57 of 71 in T2t5. There were 77 individual mice (i.e., mouse ID) in our analyses from a starting total of 90 tested.

Performance Metrics

Time to Platform

The successful mice spent significantly less time to find the platform in T1t5 (16.6 ± 1.73 s; mean \pm SE) compared to their naïve performance in T1t1 (27.1 ± 2.91 s; $t(162) = 3.315$, $p = .006$; Figure 1A). Mice performed similarly in T2t1 (18.5 ± 1.78 s) compared to T1t5 ($t(153) = -0.796$, $p = .856$; Figure 1A); performance in T2t1 was significantly better than the naïve performance T1t1 ($t(163) = 2.657$, $p = .043$). The performance in T2t5 (12.8 ± 1.48 s) was not significantly better compared to T1t5 ($t(148) = 1.178$, $p = .318$; Figure 1A). Lastly, performance in T2t5 was not significantly better compared to T2t1 ($t(148) = 2.543$, $p = .058$; Figure 1A).

Figure 1*Effect of Learning and Memory on Four Response Variables*

Note. The effect of learning (T1t1 – T1t5), memory (T1t5 – T2t1), and experiment performance (T1t5 – T2t5; see text for details) on A) time to reach the platform, B) distance travelled within one trial, and C) proportion of time spent in the target quadrant (i.e., the quadrant that contained the platform). Trials used in statistical analyses are in solid circles and bold (T1t1 = Test 1 trial 1, T1t5 = Test 1 trial 5, T2t1 = Test 2 trial 1, T2t5 = Test 2 trial 5), intermediate trials are shown as open circles. Error bars indicate upper and lower 95 % confidence intervals.

Distance Travelled

Successful mice travelled significantly less in T1t5 (3.91 ± 0.37 m; mean \pm SE) compared to their naïve performance in T1t1 (7.75 ± 0.68 m; $t(160) = 5.391$, $p < .001$; Figure 1B). Although mice travelled significantly further in T2t1 (5.49 ± 0.43 m; $t(151) = -2.915$, $p = .021$) compared to T1t5 (Figure 1B), mice still travelled significantly less in T2t1 compared to the naïve trial T1t1 ($t(163) = 2.972$, $p = .018$). The performance in T2t5 (3.18 ± 0.32 s) was not significantly better compared to T1t5 ($t(147) = 1.571$, $p = .399$; Figure 1B). Lastly, performance in T2t5 was significantly better compared to T2t1 ($t(147) = 4.542$, $p < .001$; Figure 1A).

Time Spent in Target Quadrant

Successful mice did not spend a significantly higher proportion of time in the target quadrant in T1t5 (30 ± 3 %; mean \pm SE) compared to the naïve trial T1t1 (21 ± 3 %; $t = -2.290$, $df = 165$, $p = .105$; Figure 1C). There was no significant difference in proportion of time spent in target quadrant between T1t5 and T2t1 (25 ± 3 %; $t(153) = 1.404$, $p = .499$); T2t1 was not significantly different than the naïve performance T1t1 ($t(168) = -1.127$, $p = .674$). There was also no significant difference between T2t5 (26 ± 3 %) and T1t5 ($t = 0.993$, $df = 148$, $p = .754$; Figure 1C). Lastly, performance in T2t5 was not significantly better compared to T2t1 ($t(150) = -0.421$, $p = .975$; Figure 1A).

Discussion

In this study, we explored a novel application of the Morris Water Maze (MWM) by modifying it for field use to test the learning and memory of free-living white-footed mice within their natural environment. By assessing several variables that serve as standard metrics for learning and memory within the MWM, we were able to illustrate the ability of mice to successfully navigate within a modified version of an established learning and memory paradigm. Specifically, we found that mice learned across trials, as indicated by a reduction in time and distance traveled to platform from the first to fifth trial in Test 1 (T1t1 compared to T1t5), and remembered the task, as indicated by the similar performance in time to find platform after the two-hour rest period (T1t5 compared to T2t1) and significant increase in performance in both time and distance to platform in T2t1 as compared to T1t1. In our trials, performance did not further increase after the first 5 trials, as indicated by the similar performance of mice between T1t5 and T2t5 (although our experiment did not test for maximum performance levels overall). Lastly, we found no difference in quadrant use across any of the trials regardless of test and suggest that this metric not be used to estimate learning or memory in our modified field version. Overall, our results provide conclusive evidence that the modified field-version of the MWM test can be used in free-living mice to measure learning and memory.

The reduction in time and distance traveled to find the platform across trials, that we found, is one of the simplest and most robust measures of learning within the MWM (Morris, 1984). This has been found many times in laboratory studies (Baldan Ramsey & Pittenger, 2010; Barry & Commins, 2019; Morris, 1984; Vorhees & Williams, 2006). For example, Barnhart et al. (2015) found that juvenile mice decreased their time to platform over a 7-day training period. Other studies have found that environmental or toxicological stressors may impair learning and memory, by disrupting the ability of mice to reduce time to platform (D'Hooge & De Deyn, 2001).

Our results also demonstrate robust evidence that mice were able to remember the task. In laboratory studies using the MWM memory is often measured after a 5-6 day training period and a 24 hr delay (Voorhees & Williams, 2014); however, in our modified version we demonstrated that memory occurred after 5 consecutive trials within a day and a 2-hr delay. This difference in our set-up and typical lab-based studies is important as short spacing intervals between trials can lead to improved retention of the task over short-time intervals, however, longer spacing intervals may be required to test for longer-term retention (e.g., days) of the task (Smith & Scarf, 2017; Smolen et al., 2016). Given the likely importance of

longer-term retention in wild animals, future studies should examine the effects of training intervals on longer retention over a period of days. Regardless, our current results clearly show the validity of this technique for use in the field where holding wild animals may not be feasible or may disrupt the scientific validity and purpose of the study.

Beacons vs Landmarks

There can be many strategies animals use for navigation across a landscape (Shettleworth, 1998). In this study we have focused on cues external to the body, or allocentric navigation, which consists of the use of beacons and landmarks. Beacons are located close to the goal (e.g., nest, refuge, food patch) and animals learn to associate the beacon with the goal, as such only a single beacon is needed (Shettleworth, 1998). Landmarks are located distally from the goal and animals need to use multiple cues to localize themselves within the landscape to find the goal, as such multiple cues are needed (Leonard & McNaughton, 1990). Most species use both information sources when navigating, however, they may prefer one over the other (Bennett, 1993; Vlasak, 2006). Such preference may depend on the species and the environment in which they live (e.g., Brodbeck, 1994; Brodbeck & Shettleworth, 1995; Cheng, 1995). For example, Bruck and Mateo (2010) used a radial maze and found that wild Belding's ground squirrels reached the criterion faster when they were able to use both beacon and landmark cues as compared to just beacon cues (although they were also successful in their navigation with just beacon cues). Furthermore, squirrels taken from beacon-rich environments reached the criterion faster as compared to those from beacon-poor environments whether landmark cues were present or not. In our study, we also found mice successfully navigated to their goal (the platform) with the sole use of a beacon. While in our study landmark cues were present (trees in the background) mice needed to ignore such cues (opposite to Bruck and Mateo where they could successfully use such cues) and focus on the beacon given the platform with the flag was moved between trials. If mice used landmark cues they would not navigate to the appropriate location. Future studies could provide insights into wild mice spatial navigation by comparing beacon and landmark cue usage and importance, however, critical in such a comparison would be to control for the landmarks available to mice.

Use in the Field

Our modified version of the MWM allows for the assessment of learning and memory in field-based studies and supports the expansion of cognitive research beyond laboratory studies, as it eliminates the need for transport, housing, and acclimation of wild animals *ex situ*. Studies show that the transport and captivity of such animals can activate the stress response and result in periods of extended change in stress-coping mechanisms (Adams et al., 2011; Dickens et al., 2009), which ultimately could bias and produce artifacts in the cognitive performance of animals. Additionally, keeping wild animals in captivity has been shown to result in behavioral changes such as reduced activity, loss of response to predators, and alteration of defensive behaviors (O'Regan & Kitchener, 2005; Price, 1999). Alternatively, Bruck and Mateo (2010) found that juveniles (5 days post weaning) reared in the lab and those brought into the lab had similar navigational learning ability. Keeping wild animals in captivity may also provide an ability to compare different populations under constant and modified environmental condition (e.g., Dyer & Dickinson, 1994; Hurly & Healy, 1996). Yet regardless of the changed or unchanged phenotype of wild-caught captive animals, the objective of the study may be to examine learning and memory within the animal's natural habitat. It is also suggested that animals taken from the wild should be held in captivity only when absolutely necessary and durations should be restricted to those that do not affect their ability to be safely released (Soulsbury et al., 2020). By eliminating multi-day training periods that are required by traditional MWM applications, information regarding cognitive performance can be gathered even in circumstances that are time-sensitive. Importantly, allowing animals to remain at the site of capture helps provide researchers with data that reflects what animals experience in their natural environment.

In terms of the practical use of this technique, the measure of time taken to find the platform has a useful advantage of being obtained without the use of tracking software. Distance travelled before finding

the platform requires video storage and tracking software but is not affected by swimming speed (as time is). Time taken and distance travelled are often correlated variables (e.g., increased distanced travelled means greater time taken, assuming equal swimming speed). Across trials, our results showed sharp reductions in both time and distanced travelled across trials and that mice performance level had plateaued by the end of the first test (T1t5), given there was little difference in performance between T1t5, T2t1, and T2t5 across metrics. This limit of the learning curve, demonstrates that mice have found the most efficient route to the goal. In laboratory studies, this often signals the end of the training phase as the animals have optimized their performance within the maze (Bodnoff et al., 1995; Pereira & Burwell, 2015; Vorhees & Williams, 2014). However, it is important to note that this may not equate to the maximum performance level overall (as is often demonstrated in laboratory studies), given that our experiment occurred only during one day and there may be non-linear learning curves that occur across days. For practical use in the field, measuring time taken to find the platform may be sufficient and the preferred method for those who do not have the software capability of video tracking and analysis required for measuring distance travelled. However, we suggest that distance travelled may be the better metric for assessing learning and memory due to the reduction in error associated with motivation to find the platform and, thus, swimming speed differences potentially effecting time measurements.

We found no relationship between time spent in the target quadrant across the trials and do not recommend this metric for measuring learning and memory in our modified field-version of the MWM test. While our results provide robust evidence that the MWM can be adapted for use in the field, there are certain aspects of our approach that we would change for the future and others we point out here as critically important (and different from the laboratory). First, in our initial attempts of this method mice would easily escape the pool and, thus, the 0.5 m high surrounding wall is critical. Second, because there is no way to heat the water or dry mice ambient daytime temperature is important. We did not perform this test at temperatures below 22°C. Third, our tracking software had difficulty identifying aspects of the apparatus. Initially we developed an automated tracking software, however, this approach often failed to locate the pool wall, the platform, and the starting location of the mouse. Our manual version performed much better, but still at times would identify a shadow as the mouse. We suggest 1) that a cover be used (e.g., a tarp or canopy tent) over the entire set-up to reduce shadows and 2) that the platform or platform flag be visibly identifiable from the above camera (e.g., a small square placed on top of the flagpole that the camera and software can readily identify). Fourth, given the importance of beacon and landmark cues in learning spatial navigation tasks we strongly advise that researchers appreciate and control for such factors (e.g., number of trees in the background, forest vs. open testing sites) as appropriate when comparing among treatments. This may be best accomplished by including a canopy tent with walls to block landmark cue availability from the surrounding environment. Lastly, while many expensive software programs are available for such work, we were able to develop our own relatively easily (and cheaply), and our results also show that simply tracking the time to platform, which can be measured in real-time, can provide robust results.

Our field-modified MWM likely also has the similar advantage of the basic laboratory-based MWM tests in that it can be further adapted to suit the needs of ongoing research. Future studies may expand this work to include variations in the setup of the maze such as the use of distal cues to test spatial navigation (Angelucci et al., 2002; Barnhart et al., 2015). Researchers may also choose to administer probe trials to assess memory consolidation, however, these should be used sparingly as they can slow the rate of learning (Vorhees & Williams, 2006). The protocol can also be altered to assess latent learning by placing the animal on the platform prior to searching as opposed to afterward (Stewart & Morris, 1993). Many of the adaptations seen in the traditionally used laboratory version of the MWM will likely carry over successfully into the field modified version. This allows for a versatile and reproducible experiment with results that can be analyzed in a similar fashion to those found in the laboratory.

Conclusion

Our results reflect what is seen during laboratory studies providing evidence that learning has occurred despite the absence of prolonged training. We found that time searching and distance travelled

provided the most robust outcomes, as has been most often found in the laboratory, and suggest these are the two metrics measured in future work. Time searching also has the advantage that it can be recorded in the field in real time. We believe this method will provide novel insights into many aspects of ecology and evolution as it will allow for exploration of cognitive changes to both predictable and unpredictable environmental disturbances that animals face; e.g., predation risk, anthropogenic disturbance, seasonal changes (e.g., caching behavior), dispersal, etc. For example, exploring how urbanization effects learning and memory of free-living mice may provide unique insights into human impacts on wildlife. Our research strives to illustrate how a well-established method of testing cognition in laboratory studies can be successfully applied to wild animals within their natural environment. This novel application of the classic Morris Water Maze may help minimize the disparity between laboratory and field studies potentially advancing cognitive research involving free-range animals.

Acknowledgments

We thank Alyssa Giordano for invaluable help in the field.

Author Contributions: LM and MJS conceived of the study, LM carried out all data collection with help of LH, LH carried out the data analysis. All authors contributed to the writing.

Funding: Funding was provided by the University of Massachusetts Dartmouth to MJS.

Conflict of Interest: The authors declare no conflict of interest.

References

- Adamec, R. E., Blundell, J., Strasser, K., & Burton, P. (2006). Mechanisms of lasting change in anxiety induced by severe stress. In N. Sato & R. Pitman (Eds.), *Brain mechanisms and clinical implications*. Springer.
- Adamec, R. E., Walling, S., & Burton, P. (2004). Long-lasting, selective, anxiogenic effects of feline predator stress in mice. *Physiology and Behavior*, *83*, 401–410.
- Adams, N. J., Farnworth, M. J., Rickett, J., Parker, K. A., & Cockrem, J. F. (2011). Behavioural and corticosterone responses to capture and confinement of wild blackbirds (*Turdus merula*). *Applied Animal Behaviour Science* *134*(3–4), 246–255.
- Alcalá, J. A., Callejas-Aguilera, J. E., Nelson, J. B., & Rosas, J. M. (2020). Reversal training facilitates acquisition of new learning in a Morris water maze. *Learning and Behavior*, *48*(2), 208–220.
- Angelucci, M. E. M., Cesário, C., Hiroi, R. H., Rosalen, P. L., & Da Cunha, C. (2002). Effects of caffeine on learning and memory in rats tested in the Morris water maze. *Brazilian Journal of Medical and Biological Research* *35*(10), 1201–1208.
- Anisman, H., & McIntyre, D. C. (2002). Conceptual, spatial, and cue learning in the Morris water maze in fast or slow kindling rats: Attention deficit comorbidity. *Journal of Neuroscience*, *22*(17), 7809–7817.
- Baldan Ramsey, L. C., & Pittenger, C. (2010). Cued and spatial learning in the water maze: Equivalent learning in male and female mice. *Neuroscience Letters*, *483*(2), 148–151.
- Barnes, C. A. (1979). Memory deficits associated with senescence: A neurophysiological and behavioral study in the rat. *Journal of Comparative and Physiological Psychology*, *93*(1), 74–104.
- Barnhart, C. D., Yang, D., & Lein, P. J. (2015). Using the Morris water maze to assess spatial learning and memory in weanling mice. *PloS one*, *10*(4).
- Barry, D. N., & Commins, S. (2019). A novel control condition for spatial learning in the Morris water maze. *Journal of Neuroscience Methods*, *318*, 1–5.
- Bennett, A. T. D. (1993). Spatial memory in a food storing corvid. 1. Near tall landmarks are primarily used. *Journal of Comparative Physiology A*, *173*, 193–207.
- Bodnoff, S. R., Humphreys, A. G., Lehman, J. C., Diamond, D. M., Rose, G. M., & Meaney, M. J. (1995). Enduring effects of chronic corticosterone treatment on spatial learning, synaptic plasticity, and hippocampal neuropathology in young and mid-aged rats. *Journal of Neuroscience*, *15*(11), 61–69.

- Brocardo, P. S., & Gil-Mohapel, J. (2012). Studying learning and memory in animals: A look into cognitive function. *Brain Research, 1935*, 2875.
- Brodbeck, D. R. (1994). Memory for spatial and local cues: A comparison of storing and a nonstoring species. *Animal Learning and Behavior, 22*, 119–133.
- Brodbeck, D. R., & Shettleworth, S. (1995). Matching location and color of a compound stimulus: Comparison of a food-storing and a nonstoring bird species. *Journal of Experimental Psychology: Animal Behavior Processes, 21*, 64–77.
- Bruck, J. N., & Mateo, J.M. (2010). How habitat features shape ground squirrel (*Urocitellus beldingi*) navigation. *Journal of Comparative Psychology, 124*, 176–186.
- Carman, H. (2002). Proximal versus distal cue utilization in spatial navigation: The role of visual acuity? *Neurobiology of Learning and Memory, 78*(2), 332–346.
- Cauchoix, M., Hermer, E., Chaine, A. S., & Morand-Ferron, J. (2017). Cognition in the field: Comparison of reversal learning performance in captive and wild passerines. *Scientific Reports, 7*(1), 1–10.
- Cheng, K. (1995). Landmark-based spatial memory in the pigeon. In D.L. Medin (Ed.) *Psychology of learning and motivation* (pp. 1–21). (Academic Press).
- Dalm, S., Grootendorst, J., De Kloet, E. R., & Oitzl, M. S. (2000). Quantification of swim patterns in the Morris water maze. *Behavior Research Methods, Instruments, & Computers, 32*(1), 134–139.
- D’Hooge, R., & De Deyn, P. P. (2001). Application of the Morris water maze in the study of learning and memory. *Brain Research Reviews, 36*, 60–90.
- Diamond, D. M., Campbell, A. M., Park, C. R., Woodson, J. C., Conrad, C. D., Bachstetter, A. D., & Mervis, R. F. (2006). Influence of predator stress on the consolidation versus retrieval of long-term spatial memory and hippocampal spinogenesis. *Hippocampus, 16*, 571–6.
- Dickens, M. J., Earle, K. A., & Romero, L. M. (2009). Initial transference of wild birds to captivity alters stress physiology. *General and Comparative Endocrinology, 160*(1), 76–83.
- Dyer, F. C. (1987). Memory and sun compensation by honey bees. *Journal of Comparative Physiology A, 160*, 621–633.
- Dyer, F. C., & Dickinson, J. A. (1994). Development of sun compensation by honey bees: How partially experienced bees estimate the sun’s course. *Proceedings of the National Academy of Science, 91*, 4471–4474.
- Fukuda, M. T. H., Françolin-Silva, A. L., & Almeida, S. S. (2002). Early postnatal protein malnutrition affects learning and memory in the distal but not in the proximal cue version of the Morris water maze. *Behavioural Brain Research, 133*(2), 271–277.
- Handley S. L., & Mithani S. (1984) Effects of alpha-adrenoceptor agonists and antagonists in a maze-exploration model of ‘fear’-motivated behaviour. *Naunyn-Schmiedeberg’s Arch Pharmacology, 327*, 1–5.
- Hébert, M., Bulla, J., Vivien, D., & Agin, V. (2017). Are distal and proximal visual cues equally important during spatial learning in mice? A pilot study of overshadowing in the spatial domain. *Frontiers in Behavioral Neuroscience, 11*, 109.
- Hoh, T., Beiko, J., Boon, F., Weiss, S., & Cain, D. P. (1999). Complex behavioral strategy and reversal learning in the water maze without NMDA receptor-dependent long-term potentiation. *Journal of Neuroscience, 19*(10), RC2–RC2.
- Hurly, T. A., & Healy, S. D. (1996). Memory for flowers in rufous hummingbirds: Location or local visual cues? *Animal Behaviour, 51*, 1149–1157.
- Krebs, C. J., Boonstra, R., Gilbert, S., Reid, D., Kenney, A. J., & Hofer, E. J. (2011). Density estimation for small mammals from livetrapping grids: Rodents in northern Canada. *Journal of Mammalogy, 92*(5), 974–981.
- Leonard, B., & McNaughton, B. L. (1990). Spatial representation in the rat: Conceptual, behavioral, and neurophysiological perspectives. In R. P. Kesner & D. S. Olton (Eds.), *Neurobiology of comparative cognition* (pp. 363–422). Erlbaum.
- Mastrangelo, M. E., Schleich, C. E., & Zenuto, R. R. (2009). Short-term effects of an acute exposure to predatory cues on the spatial working and reference memory performance in a subterranean rodent. *Animal Behavior, 77*(3), 685–692.
- Morand-Ferron J. (2017). Why learn? The adaptive value of associative learning in wild populations. *Current Opinion in Behavioral Sciences, 16*, 73–79.
- Morand-Ferron, J., Cole, E. F., & Quinn, J. L. (2016). Studying the evolutionary ecology of cognition in the wild: A review of practical and conceptual challenges. *Biological Reviews, 91*(2), 367–389.
- Morris, R. (1984). Developments of a water-maze procedure for studying spatial learning in the rat. *Journal of Neuroscience Methods, 11*(1), 47–60.

- O'Regan, H. J., & Kitchener, A. C. (2005). The effects of captivity on the morphology of captive, domesticated and feral mammals. *Mammal Review*, 35, 215–230.
- Olton, D. S., & Samuelson, R. J. (1976). Remembrance of places passed: Spatial memory in rats. *Journal of Experimental Psychology: Animal Behavior Processes*, 2(2), 97–116.
- Park, C. R., Zoladz, P. R., Conrad, C. D., Fleshner, M., & Diamond, D. M. (2008). Acute predator stress impairs the consolidation and retrieval of hippocampus-dependent memory in male and female rats. *Learning and Memory*, 15(4), 271–280.
- Pereira, I. T., & Burwell, R. D. (2015). Using the spatial learning index to evaluate performance on the water maze. *Behavioral Neuroscience*, 129(4), 533–539.
- Price, E. O. (1999). Behavioural development in animals undergoing domestication. *Applied Animal Behaviour Science*, 65, 245–271.
- Shen, Y., Specht, S. M., & Li, R. (1994). The hippocampus: A biological model for studying learning and memory. *Progress in Neurobiology*, 44(5), 485–496.
- Shettleworth, S. (1998). *Cognition, evolution and behavior* (1st ed.). Oxford University Press.
- Smith, C. D., & Scarf, D. (2017). Spacing repetitions over long timescales: A review and a reconsolidation explanation. *Frontiers in Psychology*, 8, 962.
- Smolen, P., Zhang, Y., & Byrne, J. H. (2016). The right time to learn: Mechanisms and optimization of spaced learning. *Nature Reviews Neuroscience*, 17, 77–88.
- Soulsbury, C. D., Gray, H. E., Smith, L. M., Braithwaite, V., Cotter, S. C., Elwood, R. W., Wilkinson, A., & Collins, L. M. (2020). The welfare and ethics of research involving wild animals: A primer. *Methods in Ecology and Evolution*, 11(10), 1164–1181.
- Squire, L. R. (1992). Declarative and nondeclarative memory: Multiple brain systems supporting learning and memory. *Journal of Cognitive Neuroscience*, 4(3), 232–243.
- Stewart, C. A., & Morris, R. G. M. (1993). The watermaze. In A. Sahgal (Ed.), *Behavioural neuroscience, Volume I, A practical approach* (pp. 107–122). IRL Press at Oxford University Press.
- LuNežič, A., Vojří, T., Čehovin Zajc, L., Matas, J., & Kristan, M. (2018). Discriminative correlation filter TracNer with channel and spatial reliability. *International Journal of Computer Vision*, 126(7), 671–688.
- Upchurch, M., & Wehner, J. M. (1988). Differences between inbred strains of mice in Morris water maze performance. *Behavior Genetics*, 18(1), 55–68.
- Vlasak, A. N. (2006). Global and local spatial landmarks: Their role during foraging by Columbian ground squirrels (*Spermophilus columbianus*). *Animal Cognition*, 9, 71–80.
- Vollmer-Conna, U. S., & Lemon, J. (1998). Spatial configuration and proximal cues. *Learning and Motivation* 29(1), 102–111.
- Vorhees, C. V., & Williams, M. T. (2014). Assessing spatial learning and memory in rodents. *ILAR Journal*, 55(2), 310–332.
- Vorhees, C.V., & Williams, M. T. (2006). Morris water maze: Procedures for assessing spatial and related forms of learning and memory. *Nature Protocol*, 1, 848–858.