Published in final edited form as:

Behav Brain Res. 2009 January 23; 196(2): 220–227. doi:10.1016/j.bbr.2008.09.002.

Validation of a 2-day water maze protocol in mice

Maria Gulinello^{a,*}, Michael Gertner^b, Guadalupe Mendoza^b, Brian P. Schoenfeld^c, Salvatore Oddo^d, Frank LaFerla^d, Catherine H. Choi^e, Sean M.J. McBride^c, and Donald S. Faber^b

^aBehavioral Core Facility, Department of Neuroscience, Albert Einstein College of Medicine, Bronx, NY 10461, USA

^bDepartment of Neuroscience, Albert Einstein College of Medicine, Bronx, NY 10461, USA

^cSection of Molecular Cardiology, Departments of Medicine and Molecular Pharmacology, Albert Einstein College of Medicine, Bronx, New York 10461, USA

^dDepartment of Neurobiology and Behavior, University of California, Irvine, Irvine, CA 92697, USA

^eMD-PHD Program, Department of Pharmacology and Physiology, Drexel University College of Medicine, Philadelphia, PA 19102, USA

Abstract

We present a 2-day water maze protocol that addresses some of potential confounds present in the water maze when using the aged subjects typical of studies of neurodegenerative disorders, such as Alzheimer's disease. This protocol is based on an initial series of training trials with a visible platform, followed by a memory test with a hidden platform 24 h later. We validated this procedure using aged (15–18 m) mice expressing three Alzheimer's disease-related transgenes, PS1(M146 V), APP(Swe), and tau(P301L). We also tested these triple transgenic mice (3xTG) and age and sex-matched wild-type (WT) in a behavioral battery consisting of tests of motor coordination (balance beam), spatial memory (object displacement task) visual acuity (novel object recognition task) and locomotor activity (open field). 3xTG mice had significantly longer escape latencies in the memory trial of the 2-day water maze test than WT and than their own baseline performance in the last visible platform trial. In addition, this protocol had improved sensitivity compared to a typical probe trial, since no significant differences between genotypes were evident in a probe trial conducted 24 h after the final training trial. The 2-day procedure also resulted in good reliability between cohorts, and controlled for non-cognitive factors that can confound water maze assessments of memory, such as the significantly lower locomotor activity evident in the 3xTG mice. A further benefit of this method is that large numbers of animals can be tested in a short time.

© 2008 Elsevier B.V. All rights reserved.

*Corresponding author at: Behavioral Core Facility, Department Neuroscience, Albert Einstein College of Medicine, 1410 Pelham Parkway South K912F, Bronx, NY 10461, USA. Tel.: +1 718 430 4042; fax: +1 718 430 8821. mgulin@aecom.yu.edu (M. Gulinello).

Keywords

Aging; Novel object placement task; Novel object recognition task; Spatial memory; Alzheimer's disease; Sex differences; Water maze

1. Introduction

Despite the fact that the water maze is one of the most widely used cognitive assays, there yet remain several logistical and theoretical issues that limit accurate interpretation in some instances. These issues include problems distinguishing performance from memory/learning, lack of sensitivity between treatment groups when controls perform at low levels, sex differences in apparent performance, and confounds such as anxiety and behavioral despair [1–9]. These and other issues are discussed in detail in the following sections. A particular case regards the multiple confounds likely to occur when testing aged subjects, as is essential in many rodent models of neurodegenerative diseases [6,8,10–12]. Here we present a 2-day water maze training protocol that addresses some of these issues and is suitable for assessing long-term memory in aged mice of both sexes.

Typical water maze protocols consist of a visible platform trial, hidden platform trials and a probe trial, the order of which varies [2,9,11]. The hidden platform is concealed beneath water made opaque, generally by non-toxic paint. Acquisition is typically assessed with hidden platform trials as a decreased latency to mount the platform (escape latency) and shorter swim distance across trials, although multiple other measures are also used [2,9,13]. Long-term memory is typically assessed in a probe trial in which the platform has been removed. It is widely assumed that subjects remembering where the platform was would have a shorter latency to the target quadrant (where the platform was), and a shorter latency to the target zone (the circular region where the platform was) and would spend more time in the target quadrant than predicted by chance (25%). However, in many cases, neither hidden platform nor probe trial procedures control for or assess the many factors governing escape latency, including stress and anxiety, behavioral despair and the robust natural tendency of rodents to explore novel environments, to name a few. In fact, it has also been suggested that probe trials may be accurately described as extinction procedures [12,14–16].

The alternative protocol presented here addresses some of these aforementioned issues. In this paradigm, subjects are trained using a visible platform for four trials on the first day. 24 h later they are tested in a hidden platform task. The potential advantages of initial training in a visible platform are many, including reducing variability and providing a baseline for individual performance, permitting habituation and reducing confounds of activity levels, anxiety, stress and behavioral despair. Firstly, this paradigm fosters habituation to the novel environment. Many water maze protocols include habituation trials without a platform because a large number of animals do not initially actively seek escape but continuously explore the pool. However, the absence of a platform (and the inability to control escape) may discourage further attempts to find the platform and is likely to induce a confounding stress response not conducive to habituation and subsequent performance, especially in animals sensitive to the induction of behavioral despair (indicated by floating), or anxiety (indicated by thigmotaxis) [15,17–21]. Secondly, exposing the animals to a situation in

which there is no escape, followed by one in which there is, could be considered to be a reversal learning task, which is not a spatial type of learning [14]. This would obviously not be an issue when including initial visible platform trials as a means of habituation.

A further issue with some water maze protocols is that there is no way to distinguish changes in escape latency due to learning from those due to changes in motivation to escape, anxiety, better search strategies, activity levels, etc. when the initial training trials are conducted with a hidden platform [6,12,22–26]. Changing the illumination of the room, the temperature of the water or manipulation of stress hormones is sufficient to change the apparent ability of subjects to learn and/or remember in the water maze [18–20,27–30]. The addition of a single visible platform at the end of all the trials is also unlikely to accurately assess or control for any of these variables. In fact, the data we present below suggest that a single visible platform trial initially is also not sufficient. Although many investigators assume that the visible platform trials monitor the visual acuity of the animals [31], visible platform trials are also potentially useful for many other reasons, not the least of which is to include control assessments within the existing task. Given these considerations, we designed a protocol based on initial training with a visible platform followed by long-term memory assessment 24 h later in a hidden platform trial.

We asked if the escape latency would decrease over time in the initial visible platform trials, resulting in stable, reproducible behavior from cohort to cohort, between individuals, between genotypes and between sexes. We further asked if analyzing the increased escape latency using a hidden platform trial 24 h after the final visible platform trial could sensitively assess long-term memory. We also assessed short-term memory (in mice with long-term memory deficits) by assessing the escape latencies in subsequent hidden platform trials performed 30 min apart. In order to validate this protocol for use in aged animals we used a triple transgenic murine model of Alzheimer's disease (AD). These 15–18-month-old mice express the PS1(M146 V), APP(Swe), and tau(P301L) transgenes and are referred to as 3xTG mice [32]. 3xTG mice and age- and sex-matched WT mice were tested in the novel 2-day water maze procedure presented here, in a typical water maze probe trial, and in a behavioral battery consisting of tests of motor coordination (balance beam), spatial memory (object displacement task) visual acuity (novel object recognition task) and locomotor activity (open field).

2. Methods

2.1. Subjects

Wild-type (WT, $n = 19$) and triple transgenic (3xTG, $n = 22$) mice of both sexes were 15–18-months old at the time of testing [32]. Sample sizes when divided by sex were as follows: WT female $n = 8$, male WT $n = 11$; 3xTG mice males: $n = 13$; females: $n = 9$. They were housed in groups of 3–5 with *ad lib* food and water in a 12–12 light–dark cycle. The rationale for this choice of expression, typical pathology and details of the creation and breeding of these mice are published elsewhere [32]. Briefly, 3xTG mice and controls were created as detailed in [32]. Essentially, two independent transgenes (encoding human APPSwe and human tauP301L, both under control of the mouse Thy1.2 regulatory element) were co-microinjected into single-cell embryos harvested from homozygous mutant

PS1M146 V knock-in (PS1-KI) mice. The PS1 knock-in mice were originally generated as a hybrid 129/C57BL6.

Mice were tested in two cohorts with each cohort age and sex matched. All studies were approved by the Animal Institute Ethical Committee of the Albert Einstein College of Medicine and were conducted according to NIH and IUACAAC guidelines.

2.2. Water maze

Animals were trained to criterion (90% escaping in under 60 s) in a series of visible platform trials on day 1 (D1) in a pool measuring 1.2 m in diameter and 50.8 cm deep using Viewer tracking software (Bonn, Germany). The water temperature was 24 ± 2 °C. The diameter of the platform (104 cm) was the same for both hidden and visible trials. This required four visible platform trials (V1–V4). The last visible platform trial of any animal is considered to be its post-habitation baseline, and is designated V4 (visible platform trial 4) or D1V4 (day 1, visible platform trial 4). Training of animals was staggered in time so that each mouse had the same inter-trial interval—i.e. 30 ± 10 min between each trial (for both the visible and the hidden trials).

The visible platform was outlined by a dark ring with high contrast to the white background made by the addition of white, non-toxic tempera paint. Animals failing to escape within 3 min were manually guided to the platform. All animals stayed on the platform for 5–10 s before being removed, dried and then placed in a holding cage with a heating pad to prevent hypothermia.

24 h after the last visible platform trial (D2), the animals were tested in a series of three hidden platform trials (T1–T3). As before, the trials were staggered in time to ensure a stable inter-trial interval of about 30 min and each trial was a maximum of 3 min long.

In order to compare the sensitivity and reliability of this protocol to standard training and probe trials, a subset of animals received an additional day of hidden platform trials followed 24 h later by a probe trial.

The platform remained in the same place for all of the trials. Although some protocols vary the position of the platform from trial to trial ([20,33,58]), given the age of the animals we did not utilize this variation. Animals were delivered to either the west or east quadrant (the target quadrant being designated as north) in a random manner for each trial as these are equidistant from the target. Some groups use all quadrants for entry points [34,35], but given that north quadrant entry is substantially closer and the south quadrant entry farther from the target than the other entry points, we chose not to introduce that potential source of variability into the study. High contrast visual cues were placed on the wall of the pool in each quadrant. External cues were not intentionally placed in the room, but the non-symmetrical nature of the testing room (door, sink, computer placement, etc.) possibly provided these.

All trials were recorded and analyzed using Viewer software (Biobserve, Bonn, Germany) and JMP statistical program (SAS, Cary, NC, USA).

2.3. Object placement and novel object recognition tasks

All adjunct tests began 2 weeks after completion of the water maze. The additional behavioral tests were conducted in the following order: open field, object placement, balance beam and object recognition. Comparison to a previous pilot experiment conducted on naïve mice did not indicate altered performance resulting from prior water maze experience (data not shown). The object placement test of spatial working memory [36,37] is based on the natural and robust tendency of rodents to preferentially explore novel objects. The assay was performed essentially as described previously [36,37]. Briefly, in Trial 1, mice were placed in an opaque plastic arena (16 in.²) and allowed to freely explore two identical, non-toxic objects (such as plastic, glass or ceramic items) for 6 min. High contrast visual cues were placed on the walls of the open field. The time spent exploring the objects was recorded manually with timers, after which the animal was returned to the home cage. Following a retention interval of 3 min, the animal was returned to the arena in which one of the objects was displaced in space (Trial 2). Care was taken to ensure that the intrinsic relationship between the objects was changed, and not just the position relative to a visual cue. The mouse was again allowed to explore for 3 min, during which time the exploration of both the displaced (novel) and familiar objects was scored.

In order to assess visual acuity, we employed the novel object recognition task with no retention interval between Trial 1 and Trial 2. In Trial 1, the animals were placed in the open field with two identical objects and allowed to explore for 3 min. The animals were then removed from the arena only long enough to replace one of the objects with a new object (between 30 and 60 s). In Trial 2, mice were allowed to explore again for 3 min and the amount of time exploring both the novel and familiar objects was recorded.

Data from both object recognition and object placement tasks are illustrated in the figures as a preference score, defined as (time spent exploring the displaced object/total time exploring both objects) $\times$ 100. A preference score of 50% indicates chance exploration (i.e. both novel and familiar objects are explored equally). Exploration of the objects was defined as any physical contact with an object (whisking, sniffing, rearing on or touching the object) and/or approach and obvious orientation to an object within 5 cm.

Animals were excluded from the analysis and graphs of preference score if they explored the objects for less than 3 s in either Trial 1 or Trial 2.

2.4. Open field

After allowing the animals to acclimatize to the testing room for 30 min, the animals were placed in an opaque open field (dimensions 40.5 cm²). Voluntary locomotor activity was assessed with an automated video-tracking system and Viewer software (Biobserve, Bonn, Germany) and is expressed as total tracklength (total distance traveled). The total time of the test was 3 min.

2.5. Balance beam

The latency to cross a round wooden balance beam and the number of missteps were recorded [38]. Mice were first exposed to walking over a flat wooden plank (10-cm wide) in

order to minimize latency variation in the subsequent tests and to acclimatize them to crossing over an exposed space. The starting side of the plank was highly illuminated whereas the goal side was dark and had a small plastic “hide” that contained a treat (chocolate cereal) to encourage the mice to cross the plank. Animals were also pre-exposed to the treats for 5 days prior to testing to reduce neophobia. Mice were allowed to walk across the plank one or two times until they rapidly reached the goal side. After this, animals were assessed for motor coordination using the round beam (2.5-cm diameter, 50-cm long). The latency to cross the beam was defined as the time taken to cross the beam from one side to the other. A slip was defined as any time that a paw fell under the midline of the beam.

2.6. Statistics and data analysis

Data from the water maze were analyzed in several ways. Escape latencies in the visible platform trials were analyzed in a two-way repeated measure ANOVA (genotype $\times$ sex $\times$ trial). To assess long-term memory, escape latencies in the last visible platform trial (D1V4) and the first hidden platform trial (D2T1) were compared. Firstly, a two-way repeated measure ANOVA (genotype $\times$ sex $\times$ trial) was used to assess the differences in absolute escape latencies between the final visible trials (D1V1–V4). Because the sex differences were no longer evident by the final visible platform trial, we then analyzed the memory trial without sex as a factor (genotype $\times$ trial, D1V4 and D2T1 and repeated measures). The subsequent comparison of cohort reliability was also likewise analyzed in a two-way repeated measure ANOVA (genotype $\times$ cohort $\times$ trial) without sex as a factor.

Mice with intact long-term memory capacity would be expected to rapidly go to the hidden platform in the first trial (D2T1) conducted 24 h after the final visible platform trial (D1V4). Deviations from this prediction would indicate a failure to remember the platform location. One way of assessing the divergence from the animal's post-habituation baseline trial (D1V4) would be to generate a difference score by subtracting the escape latency on D1V4 from D2T1. Alternatively, a ratio score can be generated by dividing the escape latency on D2T1 by that on D1V4. The difference score of a successful subject should be close to 0, and the ratio score about 1 (i.e. no increase in escape latency after a 24 h delay). Larger numbers in either measure indicate a longer escape latency after a 24 h delay. Both measures were analyzed using a two-way ANOVA (genotype $\times$ sex) followed by contrasts where appropriate. Statistical values for ANOVA are presented in the text of the results whereas the graphs report the contrasts. Data from the open field, balance beam and object placement tasks were likewise analyzed by ANOVA.

Data from the object placement test were analyzed in two ways. The preference score (% novel object exploration) was analyzed by ANOVA. We also analyzed the number of mice with and without a preference using a Chi-square. The criterion for preference was 53% preference score. This criterion was based on several factors one of which was the variance (i.e. how much deviation from 50% chance exploration could be expected from the variability of the dataset). Secondly, during the initial and extensive validation of the object placement tasks in the Behavioral Core Facility, it was determined that animals with higher than 53% preference scores consistently demonstrate novel object preferences when retested. Thirdly, we accounted for the error of scoring of exploration (with the same

experimenter scoring identical films) and determined that there could be 1–2% differences in preferences scores, especially at low exploration levels.

3. Results

3.1. Visible platform trials

It is generally assumed that escape is the overwhelming intention of the subjects in the water maze, and thus that differences in apparent escape abilities are stringently related to the cognitive abilities to find and remember the location of the platform. However, although it is commonly noted in the literature that animals failing to find the platform are guided to it manually, the exact numbers are rarely reported ([23,39,58]). Here we report that in the first visible platform trial (V1), the majority (63%) of the mice of both genotypes failed to voluntarily escape onto the visible platform within a typical trial length of 1 min (Fig. 1A). More than half (53%) of the mice still failed to escape within 2 min and 40% of the animals still failed to escape within 3 min. All mice generally showed decreased latency to escape over time regardless of genotype ($F_{\text{trial (1,32)}} = 29.9$; $p < 0.001$). By the last visible platform trial the majority of the subjects found the platform in less than 30 s regardless of genotype or (Fig. 1A). It is noteworthy that females and males performed similarly in the memory tests, which is illustrated in Fig. 5 and presented in detail below.

3.2. Long-term memory

Age- and sex-matched 3xTG mice performed worse than the WT 24 h after the final visible platform trial (Fig. 1A–D). Performance of each sex is illustrated individually in Fig. 5. We used several methods to compare performance on the last visible platform trial on day 1 (D1V4) to the first hidden platform trial conducted 24 h later (D2T1). Firstly, a repeated-measures ANOVA (with escape latencies on D1V4 and D1T1 as the repeated measures) demonstrated significantly longer escape latencies (Fig. 1A and B) in 3xTG compared to WT after a 24 h delay ($F_{\text{genotype} \times \text{trial (1,37)}} = 11.1$; $p < 0.01$). This strongly indicates a failure of long-term memory as all animals were performing at the same level prior to the delay (D1V4, Fig. 1A and B). In addition, the difference between the subject's performance on the first trial of day 2 and the last trial of day 1 can be generated (D2T1–D1V4; Fig. 1C). The ratio of the latency on the first trial of day 2 to that on the last trial of day 1 can also be used (D2T1/D1V4; Fig. 1D). In both these cases a larger number indicates poorer performance. While WT animals exhibit difference scores close to 0 (Fig. 1C), 3xTG animals have a significantly greater latency difference (60.8 ± 13 s; $F = 12.3$; $p < 0.01$). WT mice also have latency ratios close to 1 (Fig. 1D), indicating that they remember the location of the platform for 24 h. In contrast, 3xTG mice take an average of four times longer to escape than WT ($F = 8.9$; $p < 0.01$) despite the fact that the escape latency in the last visible platform trial was at the same level as WT. Total swim distance analysis resulted in pattern of responses similar to the latencies presented here and are thus not shown—i.e. that 3xTG mice had longer total swim distances than WT mice. In contrast swim speed was not significantly different between the genotypes (WT = 22.9 ± 0.8 and 3xTG 19.5 ± 2.3 cm/s).

3.3. Short-term memory

It may also be possible to assess short-term memory in this protocol when subjects exhibit an initial deficit in Trial 1 of day 2 (D2T1). 3xTG mice reached the same level of performance as WT by the third hidden platform trial on day 2 (Fig. 1A). The 3xTG mice also significantly reduced their escape latency across the three hidden platform trials (Fig. 1A, $F = 3.4$; $p < 0.05$). The inter-trial interval was 30 min; thus, it would seem that the 3xTG mice have relatively normal short-term memory if given sufficient trials.

3.4. Reliability

In order to assess the reliability of the protocol, two separate cohorts of animals were tested. Both WT and 3xTG ($n = 8$ – 13 for each genotype in each cohort) mice performed remarkably similarly from cohort to cohort in both the final visible platform trial and in the first hidden platform trial (Fig. 2). The deficits demonstrated by the 3xTG compared to the WT are robustly evident in both cohorts ($F_{\text{cohort A } (1,15)} = 4.2$; $p < 0.01$, $F_{\text{cohort B } (1,22)} = 3.1$; $p, 0.03$). Latency difference scores and proportions are also significantly different between WT and 3xTG for both cohorts ($p < 0.05$, for each cohort, data not shown).

3.5. Probe trial

A further advantage of this protocol is that it does not preclude further training with the hidden platform and subsequent probe trials (Fig. 3). Two further days of testing were thus conducted in a subset of animals. On day 3, the mice were trained in three hidden platform trials and on day 4 a probe trial was conducted. There was no significant difference between the 3xTG and the WT in the probe trial when memory was assessed as the latency to the target quadrant (Fig. 3A), as the % entries into the target quadrant (Fig. 3B), as the % time spent in the target quadrant (Fig. 3C), or as the latency to the target zone (Fig. 3D). Both genotypes spent more time in the target quadrant than that expected for chance performance, 25% (Fig. 3C), but given the variability, this is a modest indication of memory. These data highlight an additional problem with the probe trial measure. Usually, significant differences in target quadrant exploration are determined by comparing an experimental group to a control group. However, it is statistically difficult to ensure that either has significantly demonstrated “success” since there is no absolute baseline for a latency to reach the target quadrant, proportion of time spent in the target quadrant or proportion of entries to the target quadrant and no consistent criteria for deficits [40]. Other groups have also failed to find deficits in murine models of Alzheimer’s disease in the probe trial [41,42]. The deficits demonstrated by such subjects in the acquisition of the task are hard to interpret when sensorimotor deficits and activity differences are also evident.

3.6. Spatial memory in the object placement test

The 2-day protocol presented here has validity when compared to another test of spatial memory. The object placement task relies on the robust, innate tendency of rodents to preferentially explore objects that have been displaced in space, and which are therefore novel. In this assay, 82% of the 3xTG mice performed at chance levels, compared to 30% of the controls (Chi-square $p < 0.006$). The absolute preference score (Fig. 4) was also significantly lower in 3xTG mice compared to WT ($F_{(1,24)} = 7.6$; $p < 0.01$). It is important to

point out that a large number of mice – about 30% – had no (or low) object exploration in either Trial 1 or Trial 2 (see Section 2) and could not be included in the preference score analysis. These exclusions were equally distributed among genotypes and sexes. This is atypical for numerous cohorts of younger animals (data not shown), where virtually all subjects explore the objects reliably but has been reported for aged mice [6]. Taken together with the low locomotor activity, this lethargy is not likely to be genotype-dependent but is rather probably a characteristic of aged animals. Due to small remaining sample size, we thus included animals that had not been tested in the water maze and so do not present the correlation between performance in this assay and the water maze. No differences between naïve mice and animals with prior water maze experience and were found (data not shown).

3.7. Motor coordination, locomotor activity and novel object recognition task

Motor coordination was assessed as both the number of slips and the time taken to cross a round wooden balance beam. There were no differences between WT and 3xTG mice in the balance beam when assessed as number of slips (1.9 ± 0.4 and 1.1 ± 0.5 , respectively) or as the time it took to cross the beam (12.8 ± 1.2 s and 9.3 ± 1.2). General locomotor activity in the open field (track-length) was significantly lower in the 3xTG (471.3 ± 77 cm) than in WT (1034 ± 86 cm; $p < 0.001$).

In order to control for potential confounds of visual ability and other non-cognitive confounds, we utilized the novel object recognition assay with no delay time between the initial exposure to the two identical objects in Trial 1 and the subsequent testing with a novel object in Trial 2. Both WT ($64 \pm 10\%$) and 3xTG ($62 \pm 7\%$) preferred the novel object, indicating that they had sufficient visual acuity to distinguish between the two objects. It is important to note that although this test is often used to assess memory function, when it is employed for that purpose, substantially longer retention intervals are used than we used here (about 30 s).

3.8. Sex differences in the visible but not hidden platform trials

It is noteworthy that the females had longer escape latencies in the visible platform trials than males (Fig. 5A: $F_{\text{sex} \times \text{trial}}(3,34) = 3.7$; $p < 0.02$), but despite this initial difference, female WT ($n = 8$) performed as well as WT males ($n = 11$) in the memory trials (Fig. 5A and B). Furthermore, both sexes manifested the cognitive deficits equivalently in the 3xTG mice (males: $n = 13$; females: $n = 9$) (Fig. 5B). It should be noted that by the last visible platform trial, there were no significant differences between sexes or genotypes, thus it is also possible that females had a slower rate of habituation than males.

4. Discussion

The 2-day water maze paradigm presented here can sensitively detect phenotypic differences in long-term (24 h) spatial memory in aged animals in a murine model of Alzheimer's disease (Fig. 1) not detected by a typical probe trial (Fig. 3). This paradigm also provides good reliability between cohorts (Fig. 2) and is consistent with deficits in another test of spatial memory (Fig. 4). Furthermore, conducting several visible trials before the memory trials fosters habituation and similar baseline performance in treatment groups known to

otherwise have different performance, such as males and females (Fig. 5) or aged animals in many murine models of Alzheimer's disease [18,24,43]. Finally, this protocol provides a means to establish an individual baseline of performance (D1V4) as distinguished from acquisition or memory (Fig. 1). Thus, since the memory data can be analyzed compared to an individual's own baseline, there is increased sensitivity and reduced variability, which may be one reason that deficits between genotypes were significantly detected in the 2-day paradigm and not in the probe trial assay.

The inclusion of several initial visible platform trials has numerous advantages, one of which is that it permits the dissociation between memory-driven and non-cognitive factors that impact escape latency. The reduction in escape latency during the initial visible platform trials (Fig. 1) looks similar to learning curves often published during hidden platform trials. However, in this case, since the animals can see the platform, this decreased escape latency is unlikely to be due to spatial learning. In contrast, it is likely that the mice are habituating to a novel environment, much as they would in an open field, resulting in reduced anxiety and increased motivation to escape over the course of the visible platform trials. In fact, this hypothesis is consistent with data demonstrating an initial stress response followed by gradually diminishing stress hormone levels with repeated water maze exposures in learning rats and in non-learning yoked rats [17]. All of these factors also equivalently affect hidden platform trials when these are done without prior habituation, but it is then not possible to distinguish between a learning-induced decline in escape latency and a similar "learning curve" engendered by decreased motivation to explore, habituation, reduction in anxiety, etc. [18]. Thus, one advantage of including the visible platform trials is that these provide internal assessment and control of many factors, other than spatial memory, that also govern the speed and success rate of finding a hidden platform (or the former location of one in a probe trial) without conducting additional tests. These factors include assessment of baseline locomotor activity, individual escape latencies, rate of habituation and individual and treatment group differences in swim speed, to name just a few [4,15,18,27,44].

We considered the possibility that the animals failing to find the visible platform were visually impaired or exhibited motor deficits by conducting corroborative assays, including the balance beam, open field and novel object recognition test. Virtually all of the mice of both genotypes had sufficient visual acuity to distinguish between different objects in the novel object recognition task when the retention interval between Trial 1 and Trial 2 was less than 1 min. Furthermore, neither genotype exhibited deficits in motor coordination in the balance beam. The 3xTG did have significantly lower locomotor activity in the open field, however. While we cannot definitively rule out an influence of voluntary locomotor activity on the apparent memory capacity in the water maze, several pieces of evidence would indicate that this is not the case. Firstly, both genotypes have similar latencies to the visible platform in the final habituation trial, which would suggest that when memory is not involved, both genotypes have similar escape latencies. Furthermore, swim speed was not significantly different between the genotypes. Lastly, the 3xTG have similar escape latencies in T3 of day 2, again indicating that given sufficient trials the 3xTG have the underlying capacity to perform at control levels. This protocol thus provides several internal controls that help to differentiate between non-cognitive factors potentially affecting escape

latency that are not provided by several other water maze procedures, especially when differences in activity are evident in other assays.

Utilizing the hidden platform trial after a 24 h delay as an assessment of long-term spatial memory appears to provide several advantages over the probe trial, at least in this case. Clear evidence of learning on the part of WT mice in addition to significant differences between WT and 3xTG mice were detected after a 24 h delay in the first hidden platform trial (Fig. 1). In contrast, the probe trial revealed only modest and arguable evidence of learning on the part of the WT and no significant differences between WT and 3xTG mice (Fig. 3). There is also an objective criterion generated by this method via either the ratio score or the difference score (i.e. a latency difference of approximately 0 or a latency ratio of approximately 1), which enables a determination of the success of any treatment group in addition to permitting a comparison between treatment and control groups.

In fact, it has also been suggested that probe trials may be more accurately described as extinction procedures, and furthermore that extinction may affect interpretation of the probe trial as being indicative of spatial memory performance [14–16]. Tests such as the open space swim test [21] and the Porsolt forced swim test [45], in which the animals can swim but have no escape, are used to induce and/or assess depression in rodents.

Logistical issues regarding the probe trial may also limit its utility. A typical measure of probe trial performance is the latency to find the target zone (where the platform was). Although this intuitively seems to be a very reasonable measure, there are several practical concerns that affect this measure. Firstly, differences in apparent memory performance in this task are hard to determine if there are differences in acquisition during the typical hidden platform training trials, i.e. if all the animals have not learned the task to the same criterion. The platform itself is typically small (10–20 cm diameter) and the subject can sometimes swim almost directly to the platform location without sufficiently entering the zone (either with enough of the body or for enough time) for the tracking software to record this, making interpretation difficult. Several groups address this problem by making the target zone larger than the actual diameter of the platform and/or by decreasing the time threshold criteria for entry into the target zone. It is thus often a more equivocal measure than is generally appreciated. A further issue with the probe trial as an assessment of long-term memory is that the measures used are inherently variable and intrinsically lack power. Time spent in the quadrant zone is rarely greater than 40–60% and chance performance is 25% [46,47], thus, there is a very small window within which differences can be seen. Such issues are well known in the literature, and they have been addressed in some cases by the introduction of an “on-demand” platform, which automatically raises when the animal reaches the target zone [48], and which is used in lieu of a probe trial.

A significant disadvantage to standard water maze protocols is that they are expensive to set up and very labor intensive. This limits the number of animals that can be tested in any block and increases the time between blocks, and can thus introduce unwanted variability and potential confounds. With this protocol, 20 subjects can be run per session, 40 per week. Furthermore, escape latencies could conceivably be measured manually, which is a much

less expensive alternative that the substantial financial investment required for the software necessary to analyze the probe trials.

The lack of sex differences in long-term spatial memory evident in this protocol is noteworthy. Although it is often reported that males perform better than females in the water maze, this is by no means always the case [18,49,50]. Furthermore, it is not clear that the deficits in female performance always reflect a sex bias in spatial memory capacity [18,51]. Previous results demonstrate that the sex differences in the water maze can be abolished by familiarizing the animals with the pool [23]. Several possible explanations have empirical support. Females, who consistently have higher baseline locomotor activity [23,52], may take longer to habituate to the novel environment and/or may be more inclined to explore than to escape, or may also have differential stress responses [18,23]. Thus, the increased escape latency and longer swim distance often reported for females may, in part, reflect components of the task that are compensated with the addition of several exposures and may not be memory-related [53]. The data presented here are consistent in that females had significantly longer escape latencies in the visible platform trials but performed equal to males in the memory trials. In any case, the ability to use both males and females in studies of disease models has many advantages. This would obviously reduce the extent of breeding, permit within and between litter comparisons, and provide the capacity to accurately investigate sex differences in the course of neurodegenerative diseases and the efficacy of treatments.

Despite these advantages, the task as it is presented here is not likely to be useful for assessing short-term spatial memory unless deficits exist in long-term memory because animals with intact long-term memory perform at their own baseline after a 24 h delay, and continue to do so. However, we suggest there may be other variations of the test, which, after further validation, might be able to assess shorter-term memory. For example, although we used only a single platform placement in this instance, the hidden platform position could be changed as in other variations of the water maze protocol. Longer inter-trial intervals of >1 h could be instituted in the hidden platform trials in order to assess intermediate memory.

The protocol presented here has the advantages of having good reliability and results in the same phenotypic pattern as assessed in a non-stressful assay of spatial memory, the novel object displacement test. However, although the novel object displacement test is a very useful assay of spatial memory that does not depend on high levels of motor coordination, aversive conditioning or food or water deprivation, it does require consistent levels of object exploration and ambulation, which may be a disadvantage when using aged subjects.

There are many precedents for the protocol described here. Two-day protocols using the water escape version of the radial arm maze have been successfully employed to assess memory deficits in murine models of AD [54] and 1-day water maze protocols have also been used for other purposes [55]. Other versions of the water maze utilizing visible platforms also exist ([39,56,58]). These however, require much longer training procedures that utilized here. It has also been elsewhere reported that prior, non-spatial exposures to the water maze reduce non-cognitive confounds affecting performance ([1,57,58]). Thus, there

is no intrinsic reason why longer protocols are necessary to sensitively assess spatial memory deficits and why visible platform habituation trials could not be of benefit. However, although there are multiple versions of the water maze, many of the procedural variables have not been well validated or systematically examined. The present procedure provides a direct comparison to the probe trial and to another assay of spatial memory in addition to including a variety of other behavioral assessments in a test battery and good reliability between cohorts. These factors, in combination with the comparatively small investment of time and financial resources suggest that the 2-day water maze protocol is not only eminently suitable for the study of aged subjects, but may also have wider application of benefit to many other fields.

Acknowledgments

This research was supported by grants from the FRAXA Research Foundation to Catherine Choi and to Sean McBride.

References

1. Cain DP. Prior non-spatial pretraining eliminates sensorimotor disturbances and impairments in water maze learning caused by diazepam. *Psychopharmacol (Berl)*. 1997; 130(4):313–9.
2. D'Hooge R, De Deyn PP. Applications of the Morris water maze in the study of learning and memory. *Brain Res Brain Res Rev*. 2001; 36(1):60–90. [PubMed: 11516773]
3. Devan BD, McDonald RJ. A cautionary note on interpreting the effects of partial reinforcement on place learning performance in the water maze. *Behav Brain Res*. 2001; 119(2):213–6. [PubMed: 11165337]
4. Elias MF, Eleftheriou BE. Separating measures of learning and incentive in water maze experiments with aging inbred strains of mice. *Psychol Rep*. 1975; 36(2):343–51. [PubMed: 1144577]
5. Gerlai R. Behavioral tests of hippocampal function: simple paradigms complex problems. *Behav Brain Res*. 2001; 125(1–2):269–77. [PubMed: 11682118]
6. Luparini MR, Del Vecchio A, Barillari G, Magnani M, Prosdociimi M. Cognitive impairment in old rats: a comparison of object displacement, object recognition and water maze. *Aging (Milano)*. 2000; 12(4):264–73. [PubMed: 11073345]
7. Shirakawa K, Ichitani Y. Prolonged initiation latency in Morris water maze learning in rats with ibotenic acid lesions to medial striatum: effects of systemic and intranigral muscimol administration. *Brain Res*. 2004; 1030(2):193–200. [PubMed: 15571669]
8. van der Staay FJ. Assessment of age-associated cognitive deficits in rats: a tricky business. *Neurosci Biobehav Rev*. 2002; 26(7):753–9. [PubMed: 12470686]
9. Vorhees CV, Williams MT. Morris water maze: procedures for assessing spatial and related forms of learning and memory. *Nat Protoc*. 2006; 1(2):848–58. [PubMed: 17406317]
10. Adams J, Jones SM. Age differences in water maze performance and swimming behavior in the rat. *Physiol Behav*. 1984; 33(6):851–5. [PubMed: 6537514]
11. Brandeis R, Brandys Y, Yehuda S. The use of the Morris Water Maze in the study of memory and learning. *Int J Neurosci*. 1989; 48(1–2):29–69. [PubMed: 2684886]
12. Topic B, Dere E, Schulz D, de Souza Silva MA, Jocham G, Kart E, et al. Aged and adult rats compared in acquisition and extinction of escape from the water maze: focus on individual differences. *Behav Neurosci*. 2005; 119(1):127–44. [PubMed: 15727519]
13. Dalm S, Grootendorst J, de Kloet ER, Oitzl MS. Quantification of swim patterns in the Morris water maze. *Behav Res Methods Instrum Comput*. 2000; 32(1):134–9. [PubMed: 10758671]
14. Lattal KM, Mullen MT, Abel T. Extinction, renewal, and spontaneous recovery of a spatial preference in the water maze. *Behav Neurosci*. 2003; 117(5):1017–28. [PubMed: 14570551]

15. Schulz D, Buddenberg T, Huston JP. Extinction-induced “despair” in the water maze, exploratory behavior and fear: effects of chronic antidepressant treatment. *Neurobiol Learn Mem.* 2007; 87(4): 624–34. [PubMed: 17223365]
16. Schulz D, Huston JP, Buddenberg T, Topic B. Despair” induced by extinction trials in the water maze: relationship with measures of anxiety in aged and adult rats. *Neurobiol Learn Mem.* 2007; 87(3):309–23. [PubMed: 17079170]
17. Aguilar-Valles A, Sanchez E, de Gortari P, Balderas I, Ramirez-Amaya V, Bermudez-Rattoni F, et al. Analysis of the stress response in rats trained in the water-maze: differential expression of corticotropin-releasing hormone, CRH-R1, glucocorticoid receptors and brain-derived neurotrophic factor in limbic regions. *Neuroendocrinology.* 2005; 82(5–6):306–19. [PubMed: 16721035]
18. Clinton LK, Billings LM, Green KN, Caccamo A, Ngo J, Oddo S, et al. Age-dependent sexual dimorphism in cognition and stress response in the 3xTg-AD mice. *Neurobiol Dis.* 2007; 28(1): 76–82. [PubMed: 17659878]
19. Holscher C. Stress impairs performance in spatial water maze learning tasks. *Behav Brain Res.* 1999; 100(1–2):225–35. [PubMed: 10212070]
20. Snihur AW, Hampson E, Cain DP. Estradiol and corticosterone independently impair spatial navigation in the Morris water maze in adult female rats. *Behav Brain Res.* 2008; 187(1):56–66. [PubMed: 17913254]
21. Sun MK, Alkon DL. Induced depressive behavior impairs learning and memory in rats. *Neuroscience.* 2004; 129(1):129–39. [PubMed: 15489036]
22. Beiko J, Candusso L, Cain DP. The effect of nonspatial water maze pretraining in rats subjected to serotonin depletion and muscarinic receptor antagonism: a detailed behavioural assessment of spatial performance. *Behav Brain Res.* 1997; 88(2):201–11. [PubMed: 9404629]
23. Beiko J, Lander R, Hampson E, Boon F, Cain DP. Contribution of sex differences in the acute stress response to sex differences in water maze performance in the rat. *Behav Brain Res.* 2004; 151(1–2):239–53. [PubMed: 15084440]
24. Hartman RE, Wozniak DF, Nardi A, Olney JW, Sartorius L, Holtzman DM. Behavioral phenotyping of GFAP-apoE3 and -apoE4 transgenic mice: apoE4 mice show profound working memory impairments in the absence of Alzheimer’s-like neuropathology. *Exp Neurol.* 2001; 170(2):326–44. [PubMed: 11476599]
25. Morris R. Developments of a water-maze procedure for studying spatial learning in the rat. *J Neurosci Methods.* 1984; 11(1):47–60. [PubMed: 6471907]
26. Wolfer DP, Stagljar-Bozicevic M, Errington ML, Lipp HP. Spatial memory and learning in transgenic mice: fact or artifact? *News Physiol Sci.* 1998; 13:118–23. [PubMed: 11390774]
27. Chapillon P, Debouzie A. BALB/c mice are not so bad in the Morris water maze. *Behav Brain Res.* 2000; 117(1–2):115–8. [PubMed: 11099764]
28. Francis DD, Zaharia MD, Shanks N, Anisman H. Stress-induced disturbances in Morris water-maze performance: interstrain variability. *Physiol Behav.* 1995; 58(1):57–65. [PubMed: 7667428]
29. Rubinow MJ, Arseneau LM, Beverly JL, Juraska JM. Effect of the estrous cycle on water maze acquisition depends on the temperature of the water. *Behav Neurosci.* 2004; 118(4):863–8. [PubMed: 15301613]
30. Sandi C, Loscertales M, Guaza C. Experience-dependent facilitating effect of corticosterone on spatial memory formation in the water maze. *Eur J Neurosci.* 1997; 9(4):637–42. [PubMed: 9153570]
31. Lindner MD, Plone MA, Schallert T, Emerich DF. Blind rats are not profoundly impaired in the reference memory Morris water maze and cannot be clearly discriminated from rats with cognitive deficits in the cued platform task. *Brain Res Cogn Brain Res.* 1997; 5(4):329–33. [PubMed: 9197520]
32. Oddo S, Caccamo A, Shepherd JD, Murphy MP, Golde TE, Kaye R, et al. Triple-transgenic model of Alzheimer’s disease with plaques and tangles: intracellular Abeta and synaptic dysfunction. *Neuron.* 2003; 39(3):409–21. [PubMed: 12895417]

33. Choi SH, Woodlee MT, Hong JJ, Schallert T. A simple modification of the water maze test to enhance daily detection of spatial memory in rats and mice. *J Neurosci Methods*. 2006; 156(1–2): 182–93. [PubMed: 16621016]
34. Majlessi N, Choopani S, Bozorgmehr T, Azizi Z. Involvement of hippocampal nitric oxide in spatial learning in the rat. *Neurobiol Learn Mem*. 2008; 90(2):413–9. [PubMed: 18508394]
35. Verbois SL, Hopkins DM, Scheff SW, Pauly JR. Chronic intermittent nicotine administration attenuates traumatic brain injury-induced cognitive dysfunction. *Neuroscience*. 2003; 119(4): 1199–208. [PubMed: 12831873]
36. Ennaceur A, Delacour J. A new one-trial test for neurobiological studies of memory in rats. I. Behavioral data. *Behav Brain Res*. 1988; 31(1):47–59. [PubMed: 3228475]
37. Ennaceur A, Meliani K. A new one-trial test for neurobiological studies of memory in rats. III. Spatial vs. non-spatial working memory. *Behav Brain Res*. 1992; 51(1):83–92. [PubMed: 1482548]
38. Stanley JL, Lincoln RJ, Brown TA, McDonald LM, Dawson GR, Reynolds DS. The mouse beam walking assay offers improved sensitivity over the mouse rotarod in determining motor coordination deficits induced by benzodiazepines. *J Psychopharmacol*. 2005; 19(3):221–7. [PubMed: 15888506]
39. Savelieva KV, Rajan I, Baker KB, Vogel P, Jarman W, Allen M, et al. Learning and memory impairment in Eph receptor A6 knockout mice. *Neurosci Lett*. 2008; 438(2):205–9. [PubMed: 18450376]
40. Lindner MD. Reliability, distribution, and validity of age-related cognitive deficits in the Morris water maze. *Neurobiol Learn Mem*. 1997; 68(3):203–20. [PubMed: 9398584]
41. Arendash GW, King DL, Gordon MN, Morgan D, Hatcher JM, Hope CE, et al. Progressive, age-related behavioral impairments in transgenic mice carrying both mutant amyloid precursor protein and presenilin-1 transgenes. *Brain Res*. 2001; 891(1–2):42–53. [PubMed: 11164808]
42. Dumont M, Strazielle C, Staufenbiel M, Lalonde R. Spatial learning and exploration of environmental stimuli in 24-month-old female APP23 transgenic mice with the Swedish mutation. *Brain Res*. 2004; 1024(1–2):113–21. [PubMed: 15451372]
43. King DL, Arendash GW. Behavioral characterization of the Tg2576 transgenic model of Alzheimer's disease through 19 months. *Physiol Behav*. 2002; 75(5):627–42. [PubMed: 12020728]
44. Galea LA, Lee TT, Kostaras X, Sidhu JA, Barr AM. High levels of estradiol impair spatial performance in the Morris water maze and increase 'depressive-like' behaviors in the female meadow vole. *Physiol Behav*. 2002; 77(2–3):217–25. [PubMed: 12419397]
45. Porsolt RD, Bertin A, Jalfre M. Behavioral despair in mice: a primary screening test for antidepressants. *Arch Int Pharmacodyn Ther*. 1977; 229(2):327–36. [PubMed: 596982]
46. Klapdor K, van der staay FJ. The Morris water-escape task in mice: strain differences and effects of intra-maze contrast and brightness. *Physiol Behav*. 1996; 60(5):1247–54. [PubMed: 8916178]
47. Moy SS, Nadler JJ, Young NB, Perez A, Holloway LP, Barbaro RP, et al. Mouse behavioral tasks relevant to autism: phenotypes of 10 inbred strains. *Behav Brain Res*. 2007; 176(1):4–20. [PubMed: 16971002]
48. Buresova O, Krekule I, Zahalka A, Bures J. On-demand platform improves accuracy of the Morris water maze procedure. *J Neurosci Methods*. 1985; 15(1):63–72. [PubMed: 4058062]
49. Roof RL, Stein DG. Gender differences in Morris water maze performance depend on task parameters. *Physiol Behav*. 1999; 68(1–2):81–6. [PubMed: 10627065]
50. Warren SG, Juraska JM. Sex differences and estropausal phase effects on water maze performance in aged rats. *Neurobiol Learn Mem*. 2000; 74(3):229–40. [PubMed: 11031129]
51. Levy LJ, Astur RS, Frick KM. Men and women differ in object memory but not performance of a virtual radial maze. *Behav Neurosci*. 2005; 119(4):853–62. [PubMed: 16187814]
52. Slob AK, Huizer T, Van der Werff ten Bosch JJ. Ontogeny of sex differences in open-field ambulation in the rat. *Physiol Behav*. 1986; 37(2):313–5. [PubMed: 3737744]
53. Thifault S, Lalonde R, Sanon N, Hamet P. Comparisons between C57BL/6J and A/J mice in motor activity and coordination, hole-poking, and spatial learning. *Brain Res Bull*. 2002; 58(2):213–8. [PubMed: 12127020]

54. Alamed J, Wilcock DM, Diamond DM, Gordon MN, Morgan D. Two-day radial-arm water maze learning and memory task; robust resolution of amyloid-related memory deficits in transgenic mice. *Nat Protoc.* 2006; 1(4):1671–9. [PubMed: 17487150]
55. Kraemer PJ, Brown RW, Baldwin SA, Scheff SW. Validation of a single-day Morris Water Maze procedure used to assess cognitive deficits associated with brain damage. *Brain Res Bull.* 1996; 39(1):17–22. [PubMed: 8846103]
56. Bizon JL, Han JS, Hudon C, Gallagher M. Effects of hippocampal cholinergic deafferentation on learning strategy selection in a visible platform version of the water maze. *Hippocampus.* 2003; 13(6):676–84. [PubMed: 12962313]
57. Nieto-Escamez FA, Sanchez-Santed F, de Bruin JP. Pretraining or previous non-spatial experience improves spatial learning in the Morris water maze of nucleus basalis lesioned rats. *Behav Brain Res.* 2004; 148(1–2):55–71. [PubMed: 14684248]
58. Zhang WN, Pothuizen HH, Feldon J, Rawlins JN. Dissociation of function within the hippocampus: effects of dorsal, ventral and complete excitotoxic hippocampal lesions on spatial navigation. *Neuroscience.* 2004; 127(2):289–300. [PubMed: 15262320]

Fig. 1.

Mice expressing three Alzheimer's disease-related transgenes (3xTG) have spatial memory deficits in the 2-day water maze protocol. Performance in the visible platform trials (A, D1V1–D1V4) and spatial memory (A–D) in the water maze assessed as absolute escape latency (A and B) or as the difference in latencies (C: first hidden platform trial - last visible platform trial) or the latency ratio (D, first hidden platform trial/last visible platform trial) when the last visible platform trial (D1V4) was performed 24 h before the first hidden platform trial (D2T1). *Indicates a significant difference 3xTG (dark bars, $n = 19$) and WT (light bars, $n = 22$, $p < 0.05$) and + indicates a significant difference ($p < 0.05$) between the escape latency of the 3xTG on D1V4 and D2T1.

Fig. 2.

Long-term spatial memory performance is reliable between different cohorts. Individual cohorts of animals performed similarly in the last visible platform trial (left panel) and when tested 24 h later in the first hidden platform trial (right panel) regardless of genotype. $n = 8-13$ in each group (cohort and genotype). *Indicates significant differences between 3xTG and WT from that cohort at $p < 0.05$.

Fig. 3.

No significant difference between genotypes in the probe trial. Despite the significant differences evident in long-term memory in the 2-day version of the water maze protocol, 3xTG (gray bars) have the same latency to the target quadrant as WT (white bars) (A), have the same % entries into the target quadrant (B), spend the same proportion of time in the target quadrant (C) do not have significant differences in the latencies to the former location of the platform (target zone, D).

Fig. 4. Spatial memory deficits between genotypes in the object placement task of spatial working memory. WT mice (light gray bars, $n = 15$) have a significant preference ($*p < 0.05$) for the displaced (novel) object whereas 3xTG mice ($n = 11$) perform at chance levels (50% preference score). Preference scores are defined as the (time spent exploring the displaced object/total time exploring both objects) $\times 100$.

Fig. 5.

Sex differences in the visible platform trials but no sex differences in spatial memory in the 2-day water maze protocol. (A) WT female (open circles) and WT male (black triangles) perform equally well in the first hidden platform trial when tested 24 h after the final visible platform trial. In contrast, there were significant differences between females ($n = 8-9$) and males ($n = 11-13$) across the visible platform trials ($*p < 0.05$). (B) Spatial deficits were equally evident in female 3xTG (gray circles) and male 3xTG (black circles). Significant differences between WT and 3xTG are indicated in the previous graphs.