Different rankings of inbred mouse strains on the Morris maze and a refined 4-arm water escape task

By: Douglas Wahlsten, Sean F. Cooper, John C. Crabbe

Made available courtesy of Elsevier: http://www.elsevier.com

Abstract:

The submerged platform or Morris water escape task is widely used to study genetic variation in spatial learning and memory, but interpretation is sometimes difficult because of wall hugging, jumping off the platform, floating or non-spatial swim strategies. We modified the task by introducing four wide arms into the circular tank and adding features that reduced, eliminated, or compensated for several competing behaviors. Three versions of the 4-arm task were evaluated in detail, and the third version yielded good results for six of eight inbred strains. Furthermore, the 4-arm task could be scored adequately without computerized video tracking. Although performance on the 4-arm task was generally superior to the Morris maze, the extent of the improvement was strain dependent. Two strains with retinal degeneration (C3H/HeJ, FVB/NJ) performed poorly on both the Morris and 4-arm mazes, whereas C57BL/6J and DBA/2J did well on both mazes. A/J performed poorly on the Morris task but became very proficient on the 4-arm maze, despite its strong tendency to hug the walls of the tank. The BALB/cByJ strain, on the other hand, exhibited the best probe trial performance on the Morris maze but was very slow in acquiring the 4-arm task. We conclude that no single task can reveal the full richness of spatially guided behavior in a wide range of mouse genotypes.

Keywords: Learning; Spatial memory; Swimming; Mouse Phenome Project; Gene–environment interaction; Thigmotaxis

Article:

1. Introduction

The submerged platform water escape task is one of the most widely used tests of genetic variation in spatial memory in mice. In some studies of targeted mutations [25,38] it is the only task employed, and in studies involving more than one test of spatial memory, it is almost always included in the battery [1,14,29,34,54]. Because the mouse cannot see the submerged platform and the maze itself is perfectly symmetrical, the animal must utilize asymmetrical extra-maze visual cues in order to become proficient. Failure in the task, however, can occur as a consequence of several non-spatial factors and does not necessarily denote a deficit in spatial memory. Latency to reach the escape platform can be strongly influenced by non-spatial motor search patterns, and reliance on spatial cues is therefore tested with no platform present. Certain treatments can impair latency reduction over trials while having no effect on search behavior on a probe trial [16,32], and a strain that shows little or no reduction in latency may nevertheless show well focused searching on probe trials [33].

Interpretation is sometimes difficult because of "devious adaptations" of certain mouse strains [60], such as hugging the walls of the circular tank (thigmotaxis) in the A/J strain [11,58]. Floating is commonly seen in the BALB/c [19] and 129 strains [64]. The BALB/cByJ strain in particular becomes very wet (water logging) after a few trials and takes considerable time to dry itself [40]. Elaborate analysis of video tracking data is needed to make sense of the complex swim paths in the circular water tank [17,63]. In a massive survey of 50,000 video tracks from several mouse strains and mutations, Wolfer et al. [65] found that wall hugging and passivity accounted for considerably more variance than the spatial memory factor.

The list of strains reported to perform poorly in the submerged platform test is quite long [13,40,52], and there is some doubt about the usefulness of this test for mice in general. The task was originally designed for use with laboratory rats [37], animals derived from a species that commonly forages in the water for food [39] and can be trained to dive and swim underwater to obtain food [26]. Wild mice, on the other hand, tend to be prey animals when in the water [35] and they are buoyant and unable to dive. Laboratory mice tend to be inferior to rats in utilizing spatial cues to master the water maze [21,23] and often adopt non-spatial strategies [47]. They learn a dry maze almost as quickly as rats but not a water maze [61,66]. Mice are highly motivated to escape water, but this fact also makes the test very stressful [12,18,19,21,53], and the animals sometimes leap from the escape platform [23].
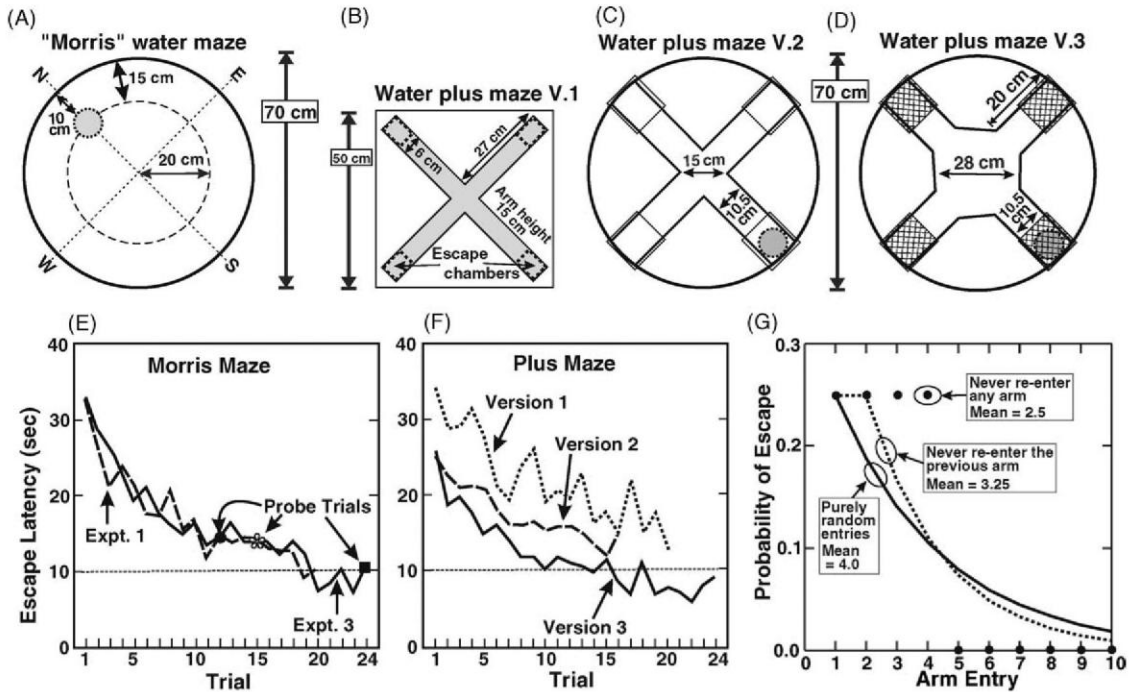

Fig. 1. Four mazes and general observations for three experiments. (A–D) Four water escape mazes. (E) Escape latencies on the Morris maze for Experiments 1 and 3. In Experiment 3, each mouse received two probe trials with no platform present. One probe was always on trial 12 (black dot), while the other probe to assess reliability was on trial 15 (asterisk) for one group of mice and trial 24 (black square) for the other group. No probe trial was given on the Morris maze in Experiment 1. (F) Escape latencies on the three versions of the 4-arm maze, computed for the eight strains that were studied with each version. (G) Probability of escaping on each trial on the 4-arm maze under different non-spatial response strategies. Derivation of these curves is described in the text.

In the experiments reported here, we asked whether alterations in the water maze apparatus or testing procedures might rescue the poor performance of some mouse strains and thereby extend the validity of the test, and whether these changes might also improve the learning of all strains. Altering the task has improved performance of mice in several other kinds of tests. For example, adding a small rim to the arms of an elevated radial maze prevents mutants with motor coordination defects from falling off the maze and allows them to demonstrate normal learning ability [24], while placing mother mice in a cone-shaped cage rescues the poor maternal care of *staggerer* mutants [28]. The problem of passive rotation in the accelerating rotarod task [31,49,58] can be effectively eliminated by using a larger diameter rod with a fine grit surface [45].

Submerged platform water escape is a complex task that is influenced by numerous processes [12,50]. Several features are clearly important for good performance. If the water is too cool, mice will rapidly become hypothermic [30] and performance is sometimes poor in cold water [43]. Performance tends to be better when the lighting of the maze is subdued [8] and handling of the animals is done gently [23]. Given the large number of potentially important variables, it was not feasible to evaluate all of them systematically. Instead, we adopted values of several parameters that are most commonly used and that yielded satisfactory results in preliminary testing.

These experiments used inbred strains to provide a survey of genetic influences. We included strains known for both good and poor performance, as well as some strains with retinal degeneration. The first experiment compared eight inbred strains on the Morris task and a relatively small water maze (Fig. 1A and B) with four arms the same width (6 cm) as is often used in the dry radial arm maze [15]. In the second experiment, 21 inbred strains were tested with a larger, improved plus water maze (Fig. 1 C) which nevertheless was plagued by strong non-spatial tendencies in many strains. Finally, eight inbred strains were compared on the Morris maze and a third version of the plus water maze (Fig. 1D) that successfully reduced these problem behaviors and yielded good learning by strains that often perform poorly in other water escape tasks.

2. Materials and methods

2.1. Mice

In Experiments 1 and 3, conducted in Edmonton only, equal numbers of male and female mice of eight inbred strains (A/J, BALB/cByJ, BTBR T+ *tf*/*tf*, C3H/HeJ, C57BL/6J, DBA/2J, FVB/NJ, 129S1/SvImJ) were shipped from the Jackson Laboratory, Bar Harbor, Maine, USA, at 6 weeks of age and tested at 9–10 weeks. All of the inbreds were on priority list A of the Mouse Phenome Project [27,41], as it was in 2000. The strain BTBR T+ *tf*/*tf* was later movedby MPP staffto priority listD (www.jax.org/phenome). Two of the strains (C3H/HeJ and FVB/NJ) suffer from retinal degeneration arising from a mutation of the retinal rod *Phosphodiesterase 6B* gene [7] (see JAX Mice Database at jaxmice.jax.org and search on the term Pde6b) and were expected to perform poorly [3,22] but were included as negative controls for the importance of visual spatial cues. Experiment 2 expanded the sample to 21 inbred strains from Jackson Labs, consisting of all strains on priority lists A and B of the MPP plus BTBR T+ *tf*/*tf*, and tested the animals simultaneously in two laboratories. Strains are identified in the results section by the strain name preceding the slash because only one substrain of any strain was studied here, except for the short A/J strain name. Mice were shipped at 7 weeks of age, tested for several behaviors at 10–12 weeks of age [44,57], and then examined for brain defects at 12.5 weeks [59]. The mice were sent in five simultaneous shipments to each lab, with each shipment usually providing two or three mice per strain per laboratory. For any one shipment to a lab, all mice of a given strain were the same sex and could have been littermates. If males of a strain were sent to one lab in a shipment, females were usually sent to the other lab. Thus, no more than four or five mice per strain could have come from the same litter, and the total sample of about 10 males and 10 females per strain must have been taken from at least five litters. The complement of strains in a single shipment was not identical for the two labs, but the net result of the five shipments was nearly equal sample sizes in the two labs.

In Experiments 1 and 3, the mice were housed four per plastic shoebox cage (Ancare 29 cm × 19 cm × 13 cm) containing 6 mm Bed-o-cob bedding and allowed free access to Edmonton tap water and Purina PMI 5001 rodent chow. The colony room was maintained at 19–21 °C and 40–60% relative humidity, and lights went on at 06:00 h and off at 18:00 h. In Experiment 2, conditions in the colony room in Portland were very similar to Edmonton in Experiments 1 and 3, except that at both sites, animals were housed two or three per cage for Experiment 2. Certain wild-derived strains (MOLF/Ei, PERA/Ei, SPRET/Ei) were maintained with filter tops on the cages in Edmonton because they came from colony rooms at the Jackson Labs that housed mice known to carry *Pasturella pneumotropica*, whereas all cages had filter tops in Portland.

2.2. Apparatus

The Morris water tank (Fig. 1A) used in both Experiments 1 and 3 was made of molded white polypropylene, 70 cm diameter and 20 cm high, and each morning it was filled with 58 L of warmed, fresh tap water to a depth of 15 cm. Water temperature was maintained at 25–26 °C throughout the day. About 60 mL of Crayola white tempera paint was added to make the water sufficiently opaque that mice could not see the submerged platform from the release point. The escape platform was round, 10 cm diameter, made from white plastic mesh mounted loosely on a clear plastic post so that the edge of the platform was 10 cm from the tank wall, and submerged 0.5 cm below the water surface. The platform could be mounted at four compass positions (N, S, E, W). Swimming was tracked with the VideoScan system from AccuScan Instruments Inc. using a Sony or Sanyo CCD TV camera having a resolution of 640 × 480 pixels, each pixel expressing one of 256 levels of gray. The 4 mm camera lens allowed a full view of the apparatus when the camera was 85 cm above the water surface. Lighting

was provided by symmetrically located tungsten bulbs sufficient to give 100 lx of incident light at the center of the maze. The camera was held in position by an aluminum frame that allowed the mouse an almost complete view of the surroundings in the room.

The plus water maze version 1 used in Experiment 1 (Fig. 1B) was built from clear plastic to have arms 6 cm wide and 27 cm long. At the end of one or more arms was a 6 cm × 7 cm escape platform placed 0.5 cm below the water surface.

Version 2 of the plus water maze (Fig. 1C) had a submerged escape platform identical to the one used in the Morris task in Experiment 1, but it was always located at the end of an arm, adjacent to the tank wall. The 3 mm thick, clear plastic arms were inserted into the same 70 cm tank used for the Morris task and extended 5 cm above the water surface so that a mouse could not reach the top edge of the wall while in the water. The end of each arm was covered by a clear plastic lid 14 cm long that made it almost impossible for the mouse to climb out of the escape chamber. Extra maze cues mounted near the maze and the video tracking systems were the same in both labs, but room cues and lighting were somewhat different in the two labs. Tracking of mice of so many different coat colors was enhanced by the addition of 2 mL Crayola blue tempera paint to the 60 mL of white paint before mixing with the water.

For Experiment 3, version 3 of the plus water maze (Fig. 1D) was the same as version 2, except that the four arms were shortened in order to create a wider central zone, and a 8 cm × 10 cm hole was cut in the lid over each end of the arm and covered with 6 mm mesh to provide better ventilation and visibility. The lighting, video tracking, extra-maze cues, water color and temperature, tank diameter, escape platform size and material, and submerged depth of the platform were the same for version 3 of the 4-arm maze and the Morris maze.

2.3. Procedures

For all mazes, the order of testing mice in each apparatus was randomized with respect to strain and balanced with respect to sex and the location of the escape platform (N, S, E, or W). All testing was conducted during the light phase of the cycle beginning at 09:00h. All animals to be tested in a session were brought from the colony room on a cart and kept in their home cages in the testing room for at least 30 min before testing. Just prior to testing, a mouse was removed from its group cage and housed singly in a cage with three clean paper towels on the floor but no food or water. The technician always wore a disposable plastic glove on the hand used to grasp the mouse's tail. Training trials used a single submerged platform in the same location every trial for a particular mouse. A trial began when the experimenter gently lowered the mouse by the tail into the water at the *center* of the apparatus, facing one of three randomly assigned compass directions; the mouse was never started facing the correct location. The mouse was started in a different, randomly assigned direction in each sequence of three consecutive trials. Video tracking began with a button push after the experimenter's hand was out of the field of view of the camera, and the experimenter moved out of the mouse's field of vision until the time for handling the animal. Trials were conducted with a 60 s time limit to reach the platform, and then 30 s between trials.

In Experiment 1, done in Edmonton only, half of the mice of each strain and sex were trained on the Morris task, and half were trained with the plus water maze version 1. Pre-training on Friday consisted of one trial of free swimming for 30 s with no platform present, one trial when the mouse was placed on a submerged platform and allowed to remain for 30 s, and two trials where the mouse was started in the center of the apparatus and allowed to swim to and climb onto any of four platforms, remaining there for 10 s. The next week, each mouse received five days of training at four trials per day. The mouse was allowed to spend 10 s on the escape platform or in the escape chamber before being returned to the holding cage.

In Experiment 2, run simultaneously in Edmonton and Portland, each mouse was given the same sequence of behavioral tests—open field activity on Monday, elevated plus maze on Tuesday, accelerating rotarod on Wednesday through Friday, then water escape training the next week. The wild-derived strains were tested at the end of each day, owing to concerns about completing testing in the event of mice escaping. Pre-training was

shortened slightly from Experiment 1 in order to fit all testing into a work day. First there was a 30 s period of free swimming with no platform present. After a 30 s intertrial interval in the holding cage, the mouse was placed on the submerged escape platform in one arm and allowed to remain for 30 s. On the third pre-training trial, swimming was confined by a clear plastic barrier to one arm that had an escape platform present at the end; the mouse was started facing the barrier and allowed 30 s to reach and climb onto the platform, where it remained for 10 s. The platform was present in the same arm of the maze that would be the correct arm during training on subsequent days. Four trials were given per day for four days with a 60 s trial limit, 10 s on the platform and 30 s between trials in the holding cage.

In Experiment 3, done in Edmonton only, two groups of mice were tested on each maze. For the first group with 12 mice per strain, pre-training for the Morris maze was the same as in Experiment 1, except that the free swimming trial with no platform present was extended to 60 s to make it the same length as the probe trial at the end of training. The zone map for the VideoScan system divided the tank into four equal, pie-shaped quadrants to determine the distribution of zone occupancy times prior to training. The second pre-training trial was 30 s on the platform, and then one trial with assisted escape onto the platform was given. The next day, training commenced and continued for five days with three trials per day. Other parameters were the same as in Experiment 1. For the Morris task, the third trial on Day 4 (Trial 12 on Friday) was a 60 s probe trial with no platform present, and the third trial on Day 5 (Trial 15 on Monday) was also a probe trial. The zone map used on probe trials had the four quadrants as well as a 10 cm circle where the correct platform was located on the regular training trials, which allowed the system to record the latency to reach the platform location even on probe trials. Time over the platform location was added to the correct quadrant time for analysis. The second group of six mice per strain received three pre-training trials on the Morris task with four platforms present on each trial, just as was done for plus maze pre-training. These animals were then trained for eight days with three trials per day and received 60 s probe trials with no platform present on trials 12 and 24.

Pre-training on the Plus maze version 3 consisted of three identical escape trials on Day 1 with a submerged platform at the end of each arm. The pre-training trial limit was 30 s and the 30 s intertrial interval was spent on the platform in the escape chamber. Training at three trials per day for five days entailed a 60 s trial limit with the 30 s intertrial interval spent on the platform. If a mouse left the platform within 30 s, it was gently directed back to the platform with the experimenter's gloved hand. The second group of mice was trained for eight days.

2.4. Video analysis
The video system recorded many episodes with no gross movement (defined by VideoScan as the center of the image moving less than 5 pixels from one frame to the next, taken at 25 frames/s) when the human observer noted clear swimming movements of the limbs. The human observer also noted periods of genuine floating, especially in the circular tank, when the mouse was coasting after a period of active swimming and therefore was seen to be moving by the video system. Thus, percentage of time with no motion, as detected by video tracking, was not a valid indicator of floating, defined as a lack of limb motion for several seconds. For this reason, the observer notes on floating were used for analysis of performance during training. Observer notes on arm entries in the plus water mazes were also considered most dependable, whereas the video tracking program was useful for estimating distance traveled and swimming speed.

3. Results
3.1. Experiment 1: Morris maze versus version 1 of the 4-arm maze
Relative to the objectives of this study, Experiment 1 was a failure from which lessons were learned, and results are described briefly. The general progress of learning in both tasks is shown in Fig. 1 E and F as reduced escape latency. On average, mice did worse on version 1 of the 4-arm maze than the Morris maze. Strain DBA improved rapidly and achieved very fast escapes in both maze configurations, whereas the A/J strain did poorly in both mazes. For the strains C3H, C57BL and 129S 1, performance was markedly superior in the circular Morris tank.

In the Morris tank, wall hugging continued to pose a problem throughout training for strains A/J and BTBR. Floating was relatively uncommon for all strains except A/J in the circular tank but surprisingly prevalent in the plus tank for A/J and 129S 1. It was apparent that some strain 129S1 mice ceased swimming when limbs touched both walls of an arm at the same time. Jumping from the submerged platform of the circular tank during training was very common in BTBR and FVB mice.

3.2. Experiment 2: Version 2 of the 4-arm maze with 21 inbred strains

The narrow arms of the 4-arm maze version 1 that rendered wall hugging innocuous also interfered with proficient swimming for many mice and increased the prevalence of floating for certain strains. Consequently, a larger plus water maze was devised (Fig. 1 C) that was identical with the Morris task in most respects, including overall size, escape platform, and stimulus surroundings. Version 2 made it impossible for mice to touch both walls of an arm at the same time and facilitated changes of direction while swimming rapidly. Consequently, escape latencies were considerably faster in version 2 than 1 of the plus maze (Fig. 1F).

We discovered after completion of the study that lighting of the maze in Portland was slightly different than in Edmonton, resulting in more difficulties tracking mice in the water in Portland and somewhat longer path lengths on trials involving equal escape latencies at the two sites. After extensive inspection of the video tracks and video taped records of many trials, we decided not to utilize the estimates of distance traveled and swimming speed from Portland. The video data from Edmonton, on the other hand, were very similar to those collected in Experiments 1 and 3.

Preliminary ANOVAs were done to assess sex differences on many variables, and only one instance of a significant ($P < 0.01$) sex difference was observed—floating during training. Consequently, data for males and females were combined for all other variables for analysis.

One surprising finding was that mice of strain C58/J were poor swimmers and quickly became soaked to the extent that they would sink (nose below water) after a few seconds of struggle to stay afloat. Any mouse that sank was quickly rescued and returned to its holding cage. A criterion was established whereby training on any day was terminated if a mouse sank on two successive trials, and any animal that had training terminated on two successive days was eliminated from the experiment. Only six of 19 C58/J mice in the two labs completed training and yielded useful data on learning.

On the 30 s free swimming trial in pre-training, strains differed substantially in swimming speed (Fig. 2A; $F=10.7$, $\omega^2 =0.50$, d.f. = 20/175 in Edmonton) and speed was closely related to the number of arms entered in the plus maze. Only 21 of 379 mice were seen to float, seven being from strain 129S 1. Thus, floating was a problem that emerged later in training (Fig. 2B) rather than an unconditioned reaction to water. On the second pre-training trial when mice were started on the platform, more than half of mice in the strains BTBR (13 of 18 mice) and SPRET (4 of 7) left the platform and entered the water within 30 s, and this behavior was also relatively common in the strains NOD (11 of 24), PERA (7 of 22) and PL (7 of 20). Leaving the platform was twice as common in Portland (23%) versus Edmonton (12%).

The general progress of learning is shown by escape latencies in Fig. 2F. Repeated measures ANOVA indicated a large strain difference (Table 1) but no significant site main effect and a borderline strain × site interaction ($P = 0.04$). Performance of strain AKR was markedly superior in Portland versus Edmonton, whereas results for many other strains were remarkably similar at the two sites, but the small strain x site interaction effect was not eliminated by removing AKR from the analysis. The rate of change in latency across trials differed significantly among strains ($P<0.00001$), with certain strains, especially those with retinal degeneration, showing little or no progress. Escape latency was substantially correlated with distance traveled before reaching the platform, partly because of strain differences in swimming speed but also because of arm entry errors. The correlation between average escape latency and average distance across all mice ranged from 0.66 to 0.79 for the four days of training in Edmonton. By the fourth and last day of training, several strains had progressed to a level of proficiency where escape occurred within 10 s and mice often made no arm entry errors (Fig. 2C). Retinal

degenerate strains commonly made many arm entry errors and took relatively long paths to reach the escape platform, although those that swam fastest (FVB, PL) often escaped reasonably quickly. Floating (Fig. 2B) was most common in five strains (129S1, A/J, AKR, BALB, SJL), and in two (129S1, A/J) it was much more common in males than females (data not shown).
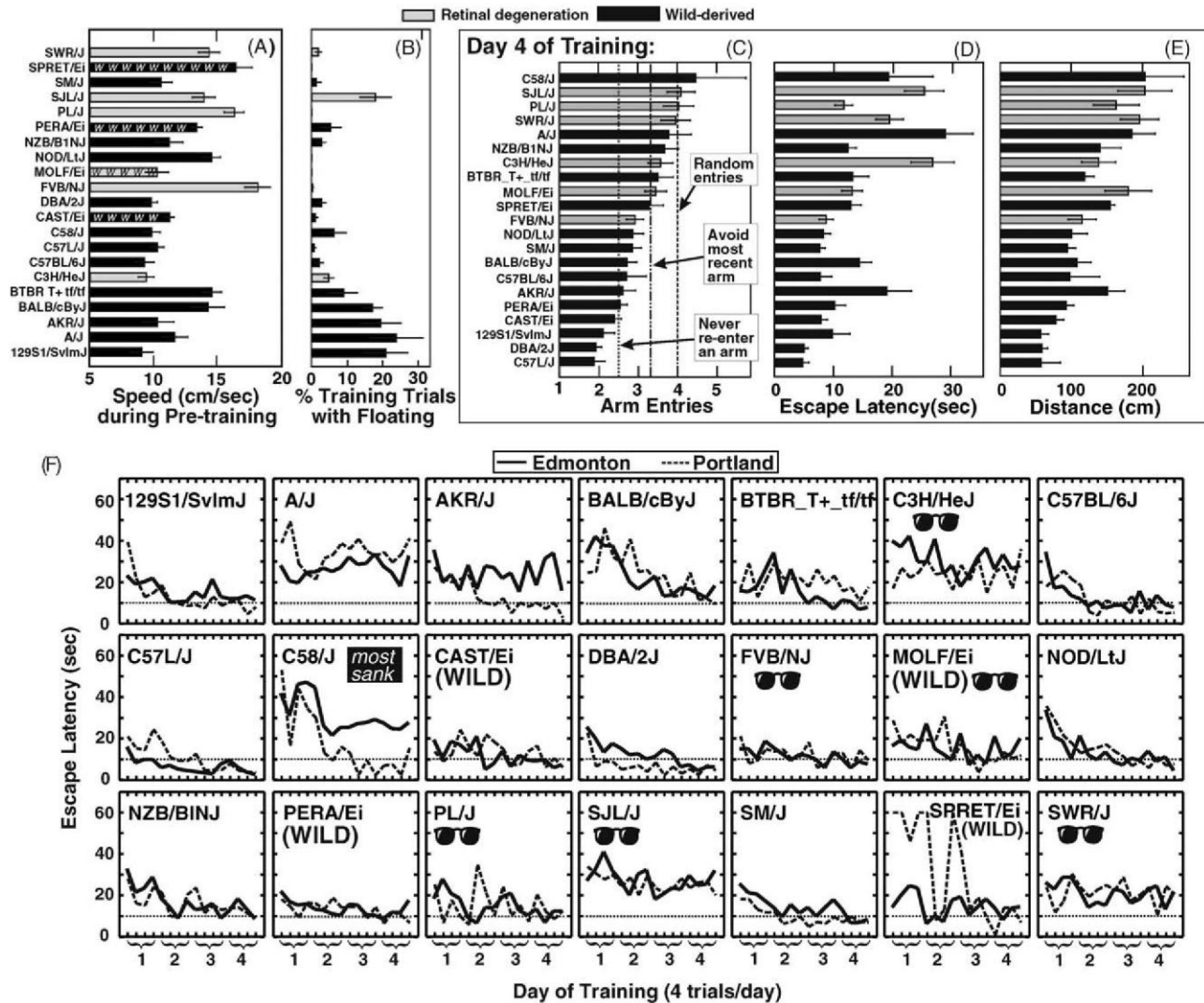

Fig. 2. Results for 21 inbred strains on version 2 of the 4-arm maze in Experiment 2 done in Edmonton and Portland. Strains with retinal degeneration are shown as grey bars and wild-derived strains are labeled w w w w. (A) Pre-training swim speed for Edmonton. (B) Percent of training trials on which the experimenter observed floating. (C) Number of arm entries on the four trials of the last day of training. Strains are ranked in C–E according to average arm entries on Day 4. Values expected from three non-spatial swim strategies shown in Fig. 1G are indicated by dashed lines. (D) Escape latencies were not closely related to number of arm entries. (E) Distance to the platform, based on data from Edmonton only, was more closely related to arm entries. (F) Escape latencies on trials run simultaneously in Edmonton and Portland. Dotted lines at 10 s indicate a rapid escape. Dark glasses denote a strain with retinal degeneration.

3.3. Patterns of arm entries in a 4-arm maze

In a symmetrical maze with A arms where only one arm has an escape platform, excellent performance is one arm entry per trial. For most mice, however, performance is less than perfect but better than chance. The sequence of arm entries provides a clue about the kind of things that are being learned and reasons for failure to master the maze. Furthermore, it allows us to predict the number of arm entries a mouse would make if it were using different strategies. We applied this analysis to the data for training Day 4 when performance was at its best.

A worst case scenario involves purely random arm entries and a stable probability of success $p = 1/A$ over several trials, such that the mouse continues to enter arms until it happens to reach the platform and escape. The probability of escape on each successive entry has a geometric distribution [20], a Pascal distribution where the first success occurs on entry k: Pr(first success on trial k)=$p(1-p)^{k-1}$. This distribution is shown in Fig. 1 G for

the case of $A = 4$ maze arms. The expected number of arm entries is $E(k) = 1/p = A$ with a variance of $A(A − 1)$. Thus, we would expect mice performing randomly to average four arm entries and make more than four arm entries on about 32% of trials.

Table 1
Analysis of indicators of learning in Experiments 2 and 3[a]

Measure	Total N	Strain effect (d.f. = 20)		Site effect (d.f. = 1)		Strain × site (d.f. = 20)	
		F	est ω^2	F	est ω^2	F	est ω^2
(A) Experiment 2 with 4-arm maze version 2 using 21 strains at two sites (Edmonton, Portland)							
Latency on 16 trials	373	12.7	0.40	0.4	0	1.6[*]	0.04
Arm entries over 16 trials	374	8.6	0.30	6.5[**]	0.02	1.8[*]	0.04
Maximum run length latency	379	13.0	0.40	4.2[*]	0.01	1.9[**]	0.05

Measure	Total N	Strain effect (d.f. = 7)		Task effect (d.f. = 1)		Strain × task (d.f. = 7)	
		F	est ω^2	F	est ω^2	F	est ω^2
(B) Experiment 3 Morris maze versus 4-arm maze version 3 in Edmonton only							
Latency on 15 trials	265	13.8	0.25	26.4	0.09	6.1	0.12
Distance on 13 trials	274	17.3	0.30	187.3	0.42	10.8	0.21
#Latencies < 10 s[b]	274	26.7	0.40	39.8	0.13	5.3	0.10
#Distances < 60 cm[b]	274	25.9	0.40	72.7	0.22	3.2[**]	0.06
Maximum run length latency[b]	274	24.1	0.38	58.0	0.18	9.9	0.19
Maximum run length distance[b]	274	13.3	0.24	23.3	0.08	2.6[**]	0.04

[a] Escape latency and distance were analysed with repeated measures ANOVA; only the between-subjects effects are shown. Distance on two probe trials on the Morris maze in Experiment 3 was not meaningful and was not included in the analysis. Data were combined for males and females. est ω^2 estimates the proportion of variance attributable to the effect in question when only that effect is compared with variation within groups.

[b] ANOVA was done on combined data for two groups receiving 15 and 24 trials. The group effect was very large but is not shown.

[*] Indicates $P < 0.05$.

[**] Indicates $P < 0.01$. All other large F ratios have $P < 0.0001$ or less.

Some mice rarely re-entered an incorrect arm that had just been explored on the previous arm entry. Whether this qualifies as genuine memory for recent location or simply a tendency not to make a sharp turn while swimming was not apparent. In either case, probability of success on the first entry is $p_1 = 1/A$, while on subsequent arm entries it will be $p_2 = 1/(A − 1)$ for random choices and the probability of escape will be Pr(first success on trial k) $= p_1$ for $k = 1$ and $(1 − p_1) p_2 (1 − p_2)^{k-2}$ for $k > 1$. The expected number of arm entries is $E(k) = (1/p_2) + p_1$ or 3.25 for $A = 4$ arms (see Fig. 1 G) and there will be more than four entries on about 22% of trials.

A mouse can also improve its chances of escape by never reentering an arm it has already entered at any time on a particular trial. This can be achieved with the aid of short term memory, but it can also result from a radial pattern of arm choices. In either case, the choice of the first arm leads to success with $p_1 = 1/A$, and the probability of success is $1/A$ for each of the first A entries but 0 for $k > A$. The expected number of arm entries is $[A(A + 1)/2]/A = 2.5$ for $A = 4$ arms (Fig. 2C) and there should never be a trial with more than four entries.

As indicated in Fig. 2C, average performance in several strains on the fourth day of training was worse than expected for animals that simply did not re-enter a previous arm, and several strains, especially ones with retinal degeneration, performed at chance levels. However, 15 of 20 strains appeared to be performing at better than chance (random entries) levels, and three strains appeared not to be using any of the three non-spatial strategies, because they averaged fewer than two arm entries.

Examination of the sequence of arm entries on all 16 trials revealed that many kinds of errors occurred. (1) A mouse sometimes swam directly into the arm towards which it was pointed when released, even though that arm was invariably wrong (start direction error). (2) A mouse sometimes entered the correct arm but then failed to swim to the end of the arm or, more commonly, reached the platform but did not climb onto it (false arm entry) and instead left the arm. (3) A mouse sometimes entered the same arm, other than the correct arm, twice or more in one trial (repeat arm entry). A special case of a repeat error was swimming from one arm to the opposite arm and then back to the first arm (back and forth). A few mice exhibited a long sequence of back and forth swimming, such as a SJL mouse that swam the sequence SNSNSNSNSNE in 55 s prior to escaping in the W arm. (4) Some animals failed to reach the platform because they ceased swimming and floated until the 60 s

limit (floating). This kind of floating differed from the float noted in Fig. 2B that sometimes occurred for only a few seconds in mice that performed reasonably well. A few animals usually entered the correct arm from a specific arm rather than going directly to the correct arm (2-arm sequence). An individual mouse could and often did commit more than one kind of error.

Table 2

Mean percent arm entry errors on trials 4–15 for 4-arm water mazes

Strain	N	Start direction error	False entry into correct arm	Repeat entry into incorrect arm	Back & forth pattern	60 s Limit with floating
Experiment 2, Plus version 2; Days 2–4						
129S1/SvImJ	20	22	10	13	5	1
A/J	18	17	**23**	**41**	13	7
AKR/J	17	32	7	15	11	8
BALB/cByJ	20	24	11	22	13	5
BTBR T+ *tf/tf*	18	**37**	**7**	**36**	27	0
C3H/HeJ*[*]	18	23	12	**37**	26	12
C57BL/6J	18	**37**	9	12	7	0
C57L/J	18	25	4	8	4	0
C58/J	6	25	15	39	19	1
CAST/Ei	18	24	4	18	6	1
DBA/2J	18	30	4	10	6	0
FVB/NJ*[*]	21	21	6	23	15	0
MOLF/Ei[*]	14	31	6	20	11	1
NOD/LtJ	24	30	10	20	13	0
NZB/B1NJ	18	23	8	29	19	0
PERA/Ei	22	31	4	15	8	3
PL/J[*]	20	20	16	**34**	**28**	0
SJL/J[*]	20	34	15	**34**	18	2
SM/J	24	**50**	7	14	5	0
SPRET/Ei	7	**63**	12	23	6	7
SWR/J[*]	20	18	15	**38**	27	0
Experiment 3, Plus version 3; Days 2–5						
129S1/SvImJ	18	16	3	10	3	0
A/J	18	17	2	19	1	3
BALB/cByJ	18	**24**	6	**25**	8	**11**
BTBR T+ *tf/tf*	17	**28**	0	16	4	0
C3H/HeJ[*]	18	**29**	2	**22**	9	2
C57BL/6J	12	17	2	8	4	1
DBA/2J	18	16	2	5	0	0
FVB/NJ[*]	18	**25**	1	**20**	7	0

Note: eight strains named in bold in Experiment 2 were also tested in Experiments 1 and 3. Strains with symbol (*) had retinal degeneration. Values shown in bold-type were particularly high.

Table 2 presents the average frequency of the various kinds of arm entry errors for the 21 strains of mice on trials 4–15. The first four trials on Day 1 of training were not included because most kinds of arm entry errors were very common for all strains at the outset of training. Start direction errors were common for most strains, but mice that learned quickly were able to suppress the tendency to swim straight ahead when released. C57BL and SM mice showed a particularly strong tendency to swim in the start direction but otherwise made few kinds of errors. Entry into the correct arm without escaping (false entry) was most common for the A/J strain that sometimes touched the platform but did not climb onto it. Repeated entry into an incorrect arm was common in retinal degenerate strains but also A/J and BTBR, and most of these strains except A/J showed many instances of back and forth swimming. Floating was common only in the C3H strain, where some mice ceased efforts to locate the platform after a few trails of rapid, unsuccessful searching. In plus maze version 2, prolonged floating did not interfere with learning; instead, it occurred after failure. An interesting pattern appeared in several mice, especially in strains NZB and SPRET, where several animals almost always entered the correct arm from a particular incorrect arm, as though they had learned to follow a specific sequence of arms and did not inhibit the tendency to enter an incorrect arm first.

Fig. 3D shows the frequency of the principal kinds of arm entry errors over days for the strains that were also tested in Experiment 3. On version 2 of the plus water maze, start direction errors were persistent for many

animals, whereas repeated arm entries gradually declined with experience. We therefore sought to alter the maze in a way that would reduce the frequency of start direction errors.

3.4. Experiment 3: Morris maze with probe trials versus version 3 of the 4-arm maze

After a series of pilot tests following Experiment 2, we adopted version 3 of the plus water maze (Fig. 1 D). This configuration, which had a larger central zone and shorter arms, made it easier for the mouse to turn and maneuver without entering an incorrect arm, while retaining the useful features of version 2 (distinct arm locations, lid over platform). We observed in Experiment 2 that mice often dove back into the water when the experimenter attempted to pick them up from the escape platform for a return to the holding cage. We therefore allowed the mouse to remain on the platform during the intertrial interval.
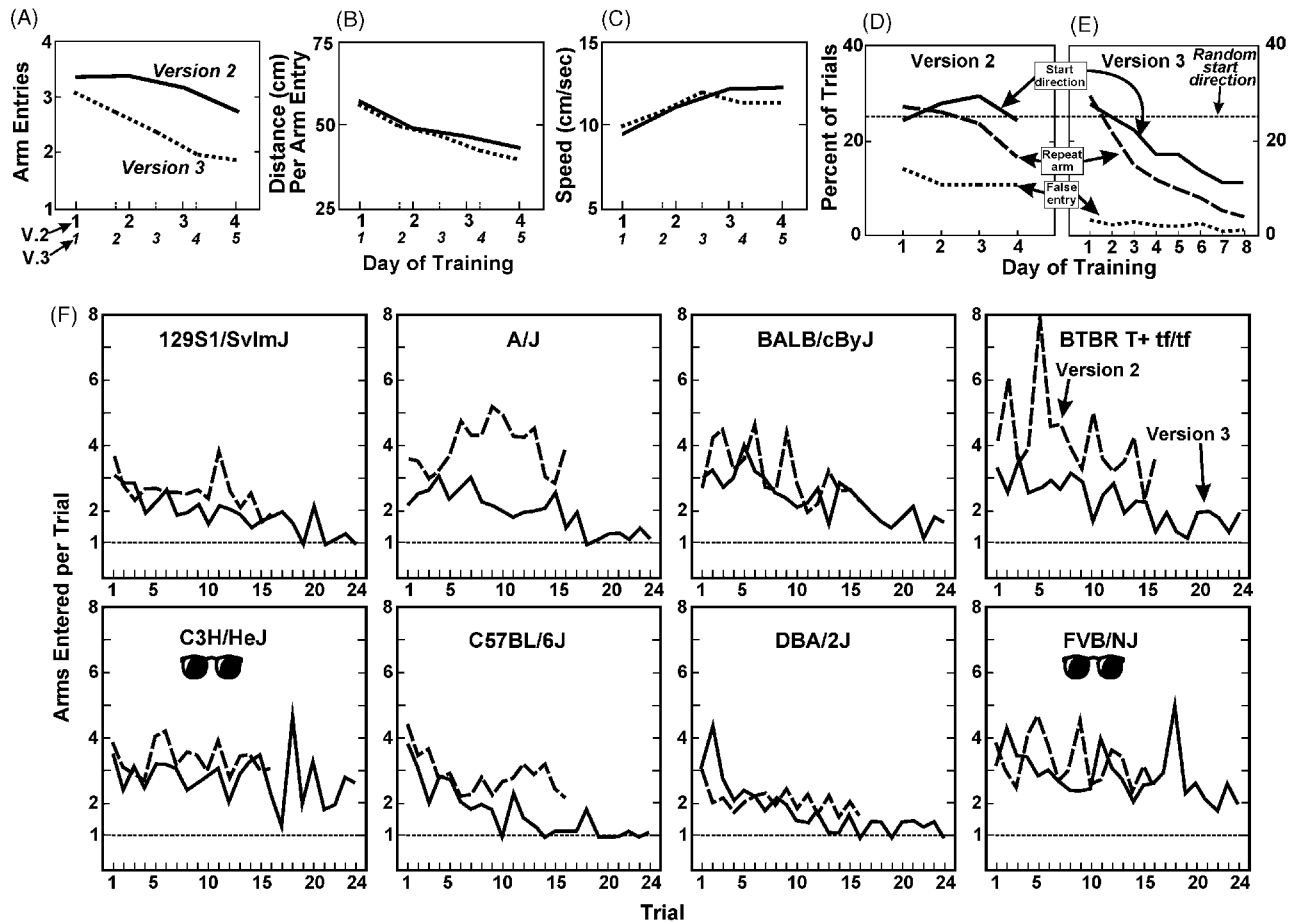

Fig. 3. Differences between plus maze versions 2 (Edmonton and Portland) and 3 (Edmonton only). (A–C) Arm entries declined more rapidly on version 3, whereas distance per entry and speed were similar. Values were computed only for the six inbred strains in both Experiments 2 and 3 that had normal eyes. (D and E) Three kinds of arm entry errors on plus maze version 2 and 3, computed for the eight strains included in both experiments. (F) Arm entries over trials for eight inbred strains run on versions 2 (Edmonton and Portland) and 3 (Edmonton) of the 4-arm water maze. Dashed line shows a perfect score, one arm entry per trial.

This change minimized handling but accelerated water logging. Consequently, we opted for only three training trials per day. A mesh opening was created in the lid to reduce condensation and enhance visibility of extra maze cues while sitting on the platform. Finally, pre-training on the 4-arm water maze was devoted entirely to escape learning, where every arm had an escape platform on three trials. This procedure was adopted after we discovered that A/J mice that did poorly on plus version 2 and even failed on visible platform escape [51,58] would learn to escape quickly when every arm had a submerged platform. By giving mice a brief experience with successful escape at the outset, we hoped to encourage active searching for the single escape platform on subsequent days. Preliminary tests indicated that three pre-training trials were sufficient for most mice to learn to escape quickly.

Sex differences were evaluated with a series of strain × sex × task ANOVAs, and the sex main effect and interactions were generally not significant ($P > 0.1$). Therefore, data were combined for males and females of

each strain. The experiment was run with three shipments of mice, and the strain C57BL was not available in the second shipment.
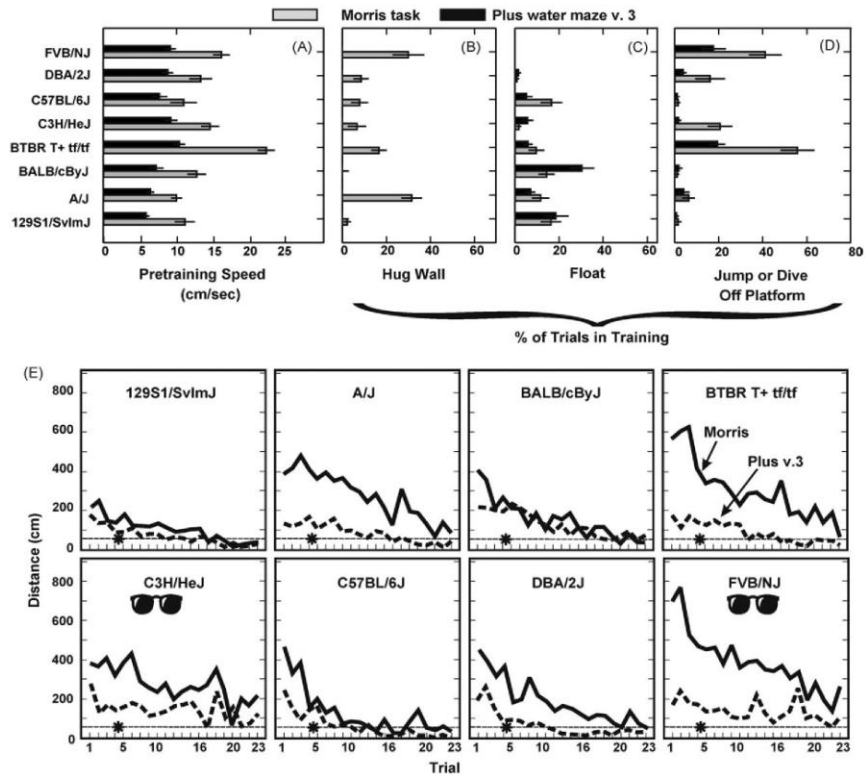

Fig. 4. Comparison of Morris maze and 4-arm maze version 3 in Experiment 3. (A) Swimming speed on the first pre-training trial. (B) Wall hugging during training seen only in the Morris maze. (C) Floating during training as observed by the experimenter. (D) Jumping from the platform or diving back into the water during training. (E) Mean escape distances. Dashed line with asterisk shows the path length (60 cm) considered to be good performance, as discussed in the text. Retinal degenerate strains (dark glasses) never achieved this level of performance on either task. Data are not shown for trials 12, 15 and 24 that were probe trials on the Morris task with no platform present.

Swimming speeds in the Morris maze (Fig. 4A) were substantially faster than in the Plus maze version 3. On the 60 s pre-training free swim, mean quadrant occupancy times in the Morris maze were 14, 16, 13 and 16 s for zones N, E, S and W, respectively, and the small zone difference was significant in a MANOVA (Wilks' lambda, $P = 0.0008$). Because the zone containing the correct platform during training was randomized across animals within a strain, the small preference could not have biased results. Most mice spent approximately 25% of the time in each zone but a few mice spent more than 35% in a zone in a 60 s trial. On the second pre-training trial on the Morris maze, several mice of the strains BTBR (9 of 12), C3H (6 of 12) and FVB (5 of 12) left the platform before the 30 s limit. For group 2 on the Morris maze, escape latencies when all four platforms were present were short but there was little reduction in latency over the three trials (Fig. 5A).

Escape learning on the Plus maze version 3 was very rapid during pre-training (Fig. 5B), and by the third trial, almost every animal entered only 1 arm and climbed onto the first platform it touched. Swimming speed increased significantly ($P < 0.000001$) from 7.8 to 10.7 cm/s over the three pre-training trials, and the rate of increase was similar for the eight strains (strain × trial interaction $F = 0.7$, $P > 0.5$), while strains differed significantly in average latency ($F = 5.2$, $P = 0.00006$; Fig. 5B). Most mice did not escape into the same arm on the three trials, and there was no indication of any preference for the first arm where a successful escape occurred.

Because the eight strains differed considerably in swimming speed (Figs. 2A and 4A), length of the path to the platform may provide a better indication of learning than escape latency. In the Plus maze version 3, entry into two or more arms always involved a path length greater than 60 cm, and therefore 60 cm was regarded as very good performance in both mazes. The general progress of learning is shown in terms of escape distance in Fig. 4E. An ANOVA (Table 1) confirmed that most strains were better on the Plus version 3 than the Morris task, but the strain x task interactions were significant and complex; several strains did much better on the Plus than the Morris task, while BALB mice were virtually the same on the two tasks. After several days of training on

Plus version 3, all strains except BALB and the retinal degenerates achieved very short escape distances. Comparing the Morris data with the same task in Experiment 1 (Fig. 1E), results were remarkably stable over a period of several months between the two experiments.
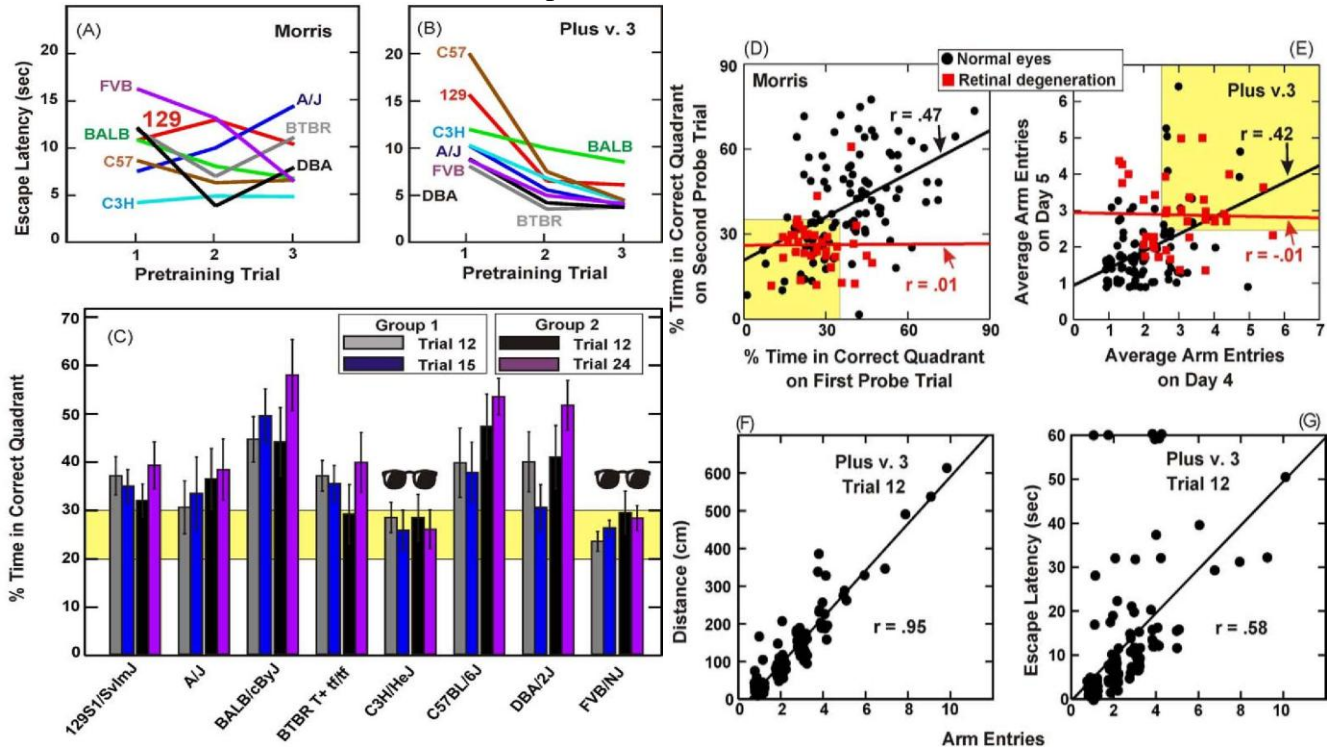

Fig. 5. *Pre-training, probe trial performance, and reliability estimates in Experiment 3. (A) Escape latencies on the Morris maze when four submerged platforms were present, one in each quadrant. (B) Escape latencies on plus version 3 with a platform in each arm improved greatly over trials for every strain but BALB. (C) Quadrant occupancy times on probe trials on the Morris maze in Experiment 3. Retinal degenerate strains (dark glasses) were in the region of random search (yellow region). This region was determined from pre-training trial 1 with no platform present, where relatively few mice spent more than 35% of the time in a quadrant in a 60 s trial with random swimming. (D) Test–retest reliability on the Morris maze was close to zero for blind mice (red), which were concentrated in the region below 35% quadrant occupancy time (yellow region). (E) Test–retest reliability on the 4-arm maze version 3 based on average number of arm entries on two successive days. (F) Correlation between arm entries and distance on trial 12 of plus maze version 3. (G) Correlation of arm entries and escape latency.*

Table 3
Indicators of learning on Morris maze (MM) and Plus maze version 3 (PLv3) in Experiment 3

Strain	No. of trials with good performance				Maximum run of good trials				Percent with four consecutive good trials			
	Latency < 10 s		Distance < 60 cm		Latency < 10 s		Distance < 60 cm		Latency < 10 s		Distance < 60 cm	
	MM	PLv3	MM	PLv3	MM	PLv3	MM	PLv3	MM	PLv3	MM	PLv3
15 Training trials												
129S1/SvImJ	8.8	9.8	7.2	7.8	5.0	6.7	4.3	3.4	67	75	67	50
A/J	3.0	8.4	2.0	5.8	1.5	5.0	0.9	2.7	8	67	0	33
BALB/cByJ	7.5	4.9	4.6	4.7	4.3	3.2	2.2	2.8	50	25	16	16
BTBR T+ *tf/tf*	8.8	9.8	3.0	4.9	5.3	5.6	1.3	2.1	67	83	8	8
C3H/HeJ	5.4	6.8	2.3	3.7	2.5	3.8	1.3	1.8	16	33	0	8
C57BL/6J	8.3	11.3	5.3	8.5	4.7	9.2	3.5	3.0	50	100	67	33
DBA/2J	8.7	11.7	3.5	6.8	5.2	9.0	1.6	3.0	91	100	11	42
FVB/NJ	5.4	8.9	1.3	3.4	2.7	4.8	0.8	1.3	16	75	0	0
24 Training trials												
129S1/SvImJ	14.0	15.2	10.8	13.0	7.5	7.2	6.0	6.3	83	83	83	100
A/J	9.0	13.0	5.7	11.3	3.3	7.5	2.3	5.7	67	100	16	67
BALB/cByJ	11.8	10.0	7.0	8.7	6.5	3.5	3.7	3.8	83	50	33	33
BTBR T+ *tf/tf*	12.3	17.0	3.3	10.8	4.2	8.8	1.5	4.2	67	100	16	67
C3H/HeJ	9.7	10.7	4.0	5.2	3.0	3.2	1.7	1.8	33	33	16	0
C57BL/6J	18.0	19.7	10.7	15.2	8.0	15.7	4.0	9.2	100	100	67	83
DBA/2J	17.5	22.2	8.0	14.7	7.3	20.7	3.3	7.7	100	100	33	67
FVB/NJ	10.0	16.2	2.8	4.8	3.5	6.7	1.5	1.8	50	100	16	0

Taking a path length of less than 60 cm or escape latency of less than 10 s as a criterion for very good performance, strain means on three indicators of learning are shown in Table 3. The number of good trials throughout training was higher on Plus version 3 than the Morris task for all strains except BALB, although the difference for 129S1 was not large. Consistency of good performance, as revealed by the longest run of consecutive good trials, was generally greater for Plus version 3 than the Morris task, except for BALB, 129S1 and the retinal degenerates. Notably long runs of short escape distances were seen with 24 training trials for strains 129S 1, A/J, C57BL and DBA. It was apparent that 15 training trials were not adequate to achieve good consistency for any strain. Four consecutive escapes with short distances would have a low probability (0.004) with random arm entries on the Plus maze, and this was taken as a criterion for unequivocally good learning. Many mice achieved this criterion within 24 training trials for escape latency on both tasks, whereas relatively few did so on the Morris task in terms of short escape distance.

Several indicators of improved performance and kinds of arm entry errors are compared for plus maze versions 2 and 3 in Fig. 3A–E. Mice achieved considerably fewer arm entries on version 3, whereas improvement in efficiency (cm per arm entry) and swimming speed were very similar for the two versions. Fig. 3D and E compare the frequency of different kinds of arm entry errors on versions 2 and 3 of the plus water maze averaged over eight strains for training trials 4 through 15. The frequency of each kind of arm entry error is shown for each strain in Table 2. Start direction errors were considerably lower for DBA and C57BL, while false arm entries were almost abolished and back and forth swimming was infrequent for all eight strains. Repeated arm entry errors almost disappeared for DBA and C57BL mice, and they were dramatically reduced for A/J, BTBR and even C3H mice. Floating was rare in plus version 3 except for BALB mice (Fig. 4C). Arm entry errors were considerably reduced over trials on version 3 for strains A/J, BTBR, C57BL/6, whereas improvement over version 2 was less pronounced for the other strains (Fig. 3). For mice receiving eight days of training on plus version 3, many in strains 129S1, A/J, C57BL and DBA achieved error free performance with good consistency.

In the Morris maze probe trials, several strains averaged considerably more than the chance level of 25% of the 60 s in the correct quadrant (Fig. 5C), especially among mice that had 24 trials. Retinal degenerate mice almost always exhibited chance levels of quadrant times, but even among mice with normal retinas there were many individuals that failed to show preference for the correct quadrant (Fig. 6). Retinal degenerate mice required greater distance and time on both probe trials to reach the platform (data not shown), and they made many more arm entry errors on the Plus maze (Figs. 3, 5 and 6).

Test–retest reliability was assessed by Pearson correlations of scores on two days. For quadrant time on the Morris maze, the correlation was about 0.5 for mice with normal retinas (Fig. 5D) but close to 0 for retinal degenerate mice. The comparable reliabilities on version 3 of the 4-arm maze were computed by averaging arm entries for three trials on one day. Test–retest reliability for the plus maze was about 0.4 for mice with normal retinas and close to 0 for retinal degenerates (Fig. 5E).

On the Morris maze, average escape latency for all mice on Days 4 and 5 was weakly correlated with average quadrant time on the two probe trials ($r = -0.26$, $t = -2.5$, d.f. = 86, $P = 0.007$ one-tail) and average distance to the platform ($r = -0.26$). For plus version 3, on the other hand, number of arm entries and distance to the platform were almost perfectly correlated (Fig. 5F), whereas arm entries and latency showed wider dispersions of scores owing to passivity by mice that failed to escape (Fig. 5G).

The single best indicator of spatial memory is time spent in the correct quadrant on probe trials on the Morris maze and arm entries on the plus maze, and we used these two criteria to rank strains (Fig. 6). Whereas retinal degenerate strains were the worst performers on both tasks, BALB mice were the best on the Morris maze but worst among sighted mice on plus version 3. For A/J mice, on the other hand, performance on the Morris maze was often poor, but on plus version 3 many became proficient, especially after eight days of training. Considering the abysmal performance of A/J on a visible platform task [51,58] as well as the Morris maze and versions 1 and 2 of the plus water escape task (Figs. 2–4), this is a noteworthy improvement.
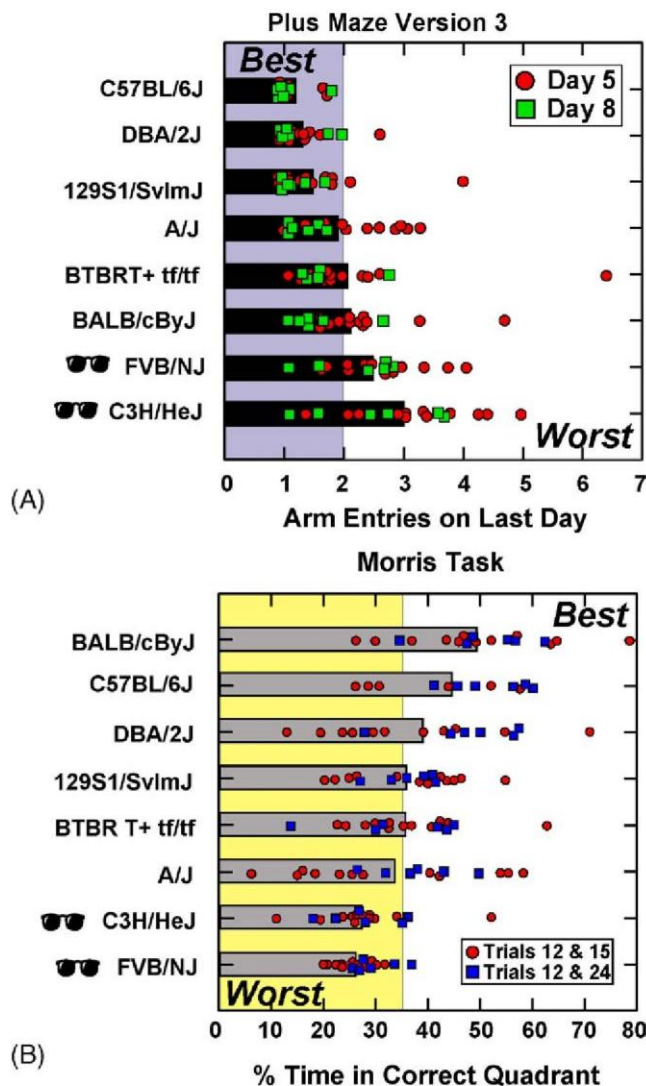
Fig. 6. Rank order of eight inbred strains on two tasks in Experiment 3. (A) Bars show strain means for average arm entries on the last day of training on plus version 3, and dots are individual scores. Dots are jittered slightly to show overlapping points. Animals with fewer than two arm entries per trial (blue region) were probably utilizing spatial relations to find the platform. Retinal degenerate mice (dark glasses) were commonly above this region, whereas for several other strains almost all mice showed excellent performance, especially after 24 training trials. (B) On the Morris maze, individuals spending less than 35% of a probe trial in the correct quadrant (yellow region) did no better than chance on probe trials. Most retinal degenerate mice were in this region, and many individuals even from strains with normal eyes showed little or no dwelling in the correct zone.

Given the different indicators of memory on the Morris and plus tasks, it is difficult to ascertain whether one yielded generally better learning than the other. On the plus maze, the strategy of never entering the same arm twice leads to an average of 2.5 arm entries (Fig. 1G), so a mouse that achieves fewer than two arm entries per trial shows evidence of utilizing spatial cues. By this criterion, almost every mouse in strains C57BL, DBA, 129S 1, A/J, BTBR and BALB showed decent if not perfect performance after eight days of training. For the Morris maze, on the other hand, random swimming usually resulted in less than 35% time in any one quadrant and most retinal degenerate mice were below 35% in the correct quadrant on probe trials (Fig. 5C and D). Taking a very generous 35% as the criterion for decent performance on probe trials, it is evident from Fig. 6B that many mice of all strains failed to show signs of utilizing spatial cues in the Morris maze. Escape latencies were also much shorter for plus version 3 than the Morris task (Fig. 1E and F), but latency by itself is not a dependable indicator of spatial memory. Combining all the evidence, we conclude that escape latencies became very short on plus version 3 for most mice because they had learned which arm contained the escape platform

(Fig. 3F). Setting aside the retinal degenerate mice, the total evidence indicated that learning and performance were on average better on plus maze version 3 than the Morris maze, but this was not true for every strain or all mice within each strain. The search patterns employed by some mice were evidently better adapted to the circular tank without arms.

4. Discussion

Version 3 of the 4-arm water escape maze greatly reduced several kinds of arm entry errors, and it yielded excellent learning for many mice of strains that did relatively poorly on the Morris maze or version 2 of the 4-arm maze, but uniformly poor performance by strains with retinal degeneration. Most remarkably, it yielded good performance for many A/J mice, a strain that failed miserably because of persistent wall hugging in every other water maze tested in our labs. It also reduced the frequency of troublesome behaviors such as jumping from the submerged platform. Thus, version 3 of the 4-arm water maze offers a convenient means for assessing memory for extra-maze cues in mice, including those that tend to hug the walls of the Morris water tank, and it may prove useful for assessing spatial learning in targeted mutations that also induce wall hugging [4].

Several principal enhancements were incorporated in version 3 of the 4-arm water escape task. (1) Wall hugging was reduced by starting mice in the center of the tank rather than adjacent to a wall, and the use of arms allowed mice to locate the escape platform despite substantial thigmotaxis while swimming. (2) Changes of direction while swimming rapidly were facilitated by using sufficiently wide arms and a large central zone at the intersection of the arms, and these features substantially reduced start direction errors. (3) Water logging and floating were reduced by using a relatively small water tank and fewer trials each day. The 4-arm maze also simplified the task by limiting the possible locations of the platform, which resulted in faster learning and therefore less recourse to floating. (4) Failure to climb on to the escape platform during training was almost eliminated by pre-training with all platforms present. (5) Jumping from the submerged platform was attenuated by placing a transparent cover over the platform to create an escape chamber and by minimizing handling between trials.

The 4-arm water maze addresses one of the most vexing issues with the Morris maze—whether to introduce a probe trial after a fixed number of training trials or after the mouse has achieved a specified criterion for good escape performance. Requiring criterion performance is logically strong but practically impossible when a wide variety of strains is studied. In the 4-arm maze, on the other hand, every trial serves as a probe trial, and the rate of progress towards consistent use of spatial cues can be monitored in all strains. Unequivocal evidence (arm entries) of the use of extra-maze spatial cues can be obtained in the 4-arm maze without the aid of an expensive video tracking system because swim path length is so closely related to number of arms entered (Pearson r = 0.95) and face validity of the arm entry count is high. Swim paths in the Morris maze during training as well as on probe trials confront the experimenter with a bewildering variety of paths that requires sophisticated computer analysis to give some sense of order to the data [17,63].

We do not argue that version 3 of the 4-arm water maze should replace the Morris maze in all situations. As shown by the highly significant and large strain x task interactions in Experiments 1 and 3 as well as the different strain rank orders in Fig. 6, no maze configuration is best for all strains. Instead, there seem to be strain-specific reactions to each kind of task [60]. Likewise, effects of aging and drugs may be specific to the kind of maze employed [9,62]. Hence, data from more than one kind of apparatus may be necessary to give an adequate understanding of spatial learning and memory [34,47]. As we have argued for motor incoordination/ataxia [10], no single test can capture adequately the genetic variation that exists in a complex behavioral domain. Testing mice in both the Morris and 4-arm tasks teaches us more about how mice navigate spatially than does either task alone, although it also requires more resources. In the future it will be interesting to study transfer of training from one kind of maze to another when they are run with identical extra-maze cues. Assessing the same strains in dry mazes is also important because of the stressful nature of swimming tasks for mice [12].

Although version 3 of the 4-arm water maze was clearly superior to the other versions reported here, it is not possible to attribute this superior performance or a specific fraction of the improvement to any one factor. Furthermore, we do not claim that the final configuration in this study is optimal. The submerged platform water escape task involves many variables that can alter results for some strains of mice [12,50]. For example, more than four water maze arms could be investigated [9]. There were simply too many factors to vary each and observe the consequences for a wide range of genetic strains. Instead, we adopted some values that gave good results and kept them the same for all apparatus in this study. For other variables, we ran pilot tests and then introduced clusters of enhancements to create a new version. For most variables, the magnitude of improved performance tends to be small, and small effects can only be detected with very large samples of mice when strain × task interactions are anticipated [55]. Combining several changes attained larger improvements.

Data from this study and our previous study of visible platform water escape [58] provide an interesting portrait of each of the eight inbred strains. Several of the strains were also tested on the Morris maze by Bolivar and Flaherty [6] and Owen et al. [40], and the similarity of our results on the Morris maze with those previous data is substantial.

The A/J strain has a very strong tendency to hug the walls of the water tank [58] and needs several experiences to learn how to mount the escape platform. Learning the 4-arm maze was slow but they did become proficient after 18 trials or six days of training. It is possible that this albino strain was impaired visually, as are many albino strains [2,3]. Most albino strains are relatively poor on the Morris maze [40], and mild visual impairment is expected to retard but not prevent acquisition of a visually guided task.

BALB/cByJ mice tended to become water logged quickly, and they should not be given more than three trials per day. It was among the best strains on the Morris maze probe trials where maze arms did not impede swimming, but it seemed to have difficulty choosing among four maze arms, especially on version 3 where its performance was highly inconsistent from trial to trial. This strain in particular may benefit from fewer trials per day over considerably more days. BALB mice are sometimes regarded as poor on the Morris maze, but a suitable set of parameters can yield decent results [8], and probe trial performance may be quite good even while latencies suggest BALB mice do not learn [33,40].

BTBR T+ *tf/tf* mice were remarkably strong swimmers for which the water held few terrors. They often jumped off the escape platform or dove back into the water without any prodding from the experimenter. These mice are not particularly wild or difficult to handle during water maze or other kinds of tests [57]. Our data suggest that their motivation to avoid water may be lower than for other strains, but it is equally plausible that they are prone to constant activity and do not remain for long in any one place. They were the only mice able to climb onto the lid of the escape compartment in the 4-arm maze. Variability between mice within this strain was very high.

C3H/HeJ are typical of what one would expect from blind mice. This strain is able to sense the general level of light in a room and entrain a light-dark cycle with the aid of retinal ganglion cells [5,42,46], but adults appear to lack pattern vision [3,7]. They swim actively but randomly on the first few trials, never execute a long series of rapid escapes, and often cease efforts to find the platform in the 4-arm maze.

C57BL/6J mice are among the best in the Morris maze, a widely recognized fact [13,48], and they also do well in versions of the 4-arm maze that do not confine their swimming. Their strong tendency to swim in the start direction is effectively combated in the version 3 maze.

DBA/2J mice are remarkable for their excellent performance in all 4-arm mazes utilized in this study, where their arm entry errors were among the lowest. Despite their rapid improvement in escape latency and distance to the Morris maze platform, many of these animals showed no preference for the correct quadrant in the probe trials. Given their superior learning of the 4-arm maze, this observation challenges the validity of the Morris

maze probe trial for this strain. Our results for DBA/2J contrast with previous observations of poor Morris maze learning for this strain [36,48,52] but are consistent with the results of others [6,40].

FVB/NJ mice are very active swimmers, and they frequently leave the escape platform and re-enter the water. Although quadrant occupancy times on Morris maze probe trials and arm entry errors in the 4-arm maze make it clear that these retinal degenerate mice lack adequate vision, they nevertheless show improved escape latencies on the 4-arm mazes. It seems likely that these mice rely on kinesthetic cues to minimize re-entry into previously visited arms, an approach that leads to a reduction but never a complete elimination of arm entry errors.

129S1/SvImJ mice are remarkable for the extreme individual variation within the strain. On every maze evaluated in this study, several mice of this strain performed very well and exhibited textbook-smooth learning curves, but invariably there were several that failed to show any improvement at all. We previously established that this variation is not related to absence of the corpus callosum in this strain [56]. The good learners are among the slowest swimmers but also show the shortest paths to the Morris maze platform and the fewest arm entry errors in the 4-arm maze. The experimenter often could sense when a 129S 1 mouse had mastered the task because it would turn in the correct direction while being held by the tail prior to release into the water, which made it difficult to point the mouse in the assigned start direction. Mice of this strain show a relatively high frequency of floating, but their episodes of floating are brief and do not substantially impair learning. It is possible that brief episodes of floating may actually enhance learning about extra-maze cues.

Expanding the sample to 21 inbred strains on version 2 of the 4-arm maze not only helped to characterize complex behavior in a wide variety of strains, a central goal of the Mouse Phenome Project, but also uncovered some unique patterns of performance in the water maze. The C58/J strain was remarkable for its tendency to sink while struggling to swim, and we suspect that the oils in their fur may be unusual. Their difficulty is specific to the water tank, whereas in several kinds of dry apparatus they are relatively active and skilled [44]. The C57L/J strain was the best performer among all 21 strains, and it showed a weaker tendency than C57BL/6J to swim straight into the opposite arm, despite its higher swim speed. The four wild-derived strains were not particularly difficult to handle during the water escape tasks, in contrast with their frequent attempts to flee during drytasks [57]. We found that they devoted a great deal of attention to grooming the wet pelage rather than plotting their next flight away from the experimenter. In the wild-derived mice, passive floating was absent.

In mouse neurobehavioral genetics there is great interest in studying mechanisms of spatial memory, but strain-specific behavioral tendencies sometimes act contrary to our purpose. Part of the process of learning spatial relations involves suppressing competing tendencies. If the apparatus and procedures make this difficult, spatial abilities may not be evident in strains where competing tendencies are strong. The 4-arm water escape maze described here reduced most competing tendencies to a low level and yielded measurable spatial learning for all strains we studied except those with rodless retinas. The task required more than 1 week of training to achieve proficiency in most strains and therefore is not ideal for high throughput screening. Further refinements especially in the visual stimuli may accelerate learning and enhance memory for spatial relations.

References
[1] Arendash GW, King DL. Intra- and intertask relationships in a behavioral test battery given to Tg2576 transgenic mice and controls. Physiol Brain Res 2002;75:643–52.
[2] Balkema GW, Drager UC. Impaired visual thresholds in hypopigmented animals. Vis Neurosci 1991;6:577–85.

[3] Balkema GW, Mangini NJ, Pinto LH, Vanable Jr JW. Visually evoked eye movements in mouse mutants and inbred strains. A screening report. Invest Ophthalmol Vis Sci 1984;25:795–800.

[4] Balschun D, Wolfer DP, Gass P, Mantamadiotis T, Welzl H, Sch¨utz G, et al. Does cAMP response element-binding protein have a pivotal role in hippocampal synaptic plasticity and hippocampus-dependent memory? J Neurosci 2003;23:6304–14.

[5] Berson DM. Strange vision: ganglion cells as circadian photoreceptors. Trends Neurosci 2003;26:314–20.

[6] Bolivar VJ, Flaherty L. Assessing autism-like behaviors in inbred strains of mice. Soc Neurosci Abstr, 2003.

[7] Chang B, Hawes NL, Hurd RE, Davisson MT, Nusinowitz S, Heckenlively JR. Retinal degeneration mutants in the mouse. Vis Res 2002;42:517–25.

[8] Chapillon P, Debouzie A. BALB/c mice are not so bad in the Morris water maze. Behav Brain Res 2000;117:115–8.

[9] Chen GH, Wang YJ, Wang XM, Zhou JN. Accelerated senescence prone mouse-8 shows early onset of deficits in spatial learning and memory in the radial six-arm water maze. Physiol Behav 2004;82:883–90.

[10] Crabbe JC, Metten P, Cameron AJ, Wahlsten D. An analysis of the genetics of alcohol intoxication in inbred mice. Neurosci Biobehav Rev 2005;28:785–802.

[11] Crabbe JC, Wahlsten D, Dudek BC. Genetics of mouse behavior: interactions with laboratory environment. Science 1999;284:1670–2.

[12] Crawley JN. What's wrong with my mouse? Behavioral phenotyping of transgenic and knockout mice. New York: Wiley-Liss; 2000.

[13] Crawley JN, Belknap JK, Collins A, Crabbe JC, Frankel W, Henderson N, et al. Behavioral phenotypes of inbred mouse strains: implications and recommmendations for molecular studies. Psychopharmacology 1997;132:107–24.

[14] Crawley JN, Paylor R. A proposed test battery and constellations of specific behavioral paradigms to investigate the behavioral phenotypes of transgenic and knockout mice. Horm Behav 1997;27:201–10.

[15] Crusio WE, Schwegler H, Brust I. Covariations between hippocampal mossy fibres and working and reference memory in spatial and non-spatial radial maze tasks in mice. Eur J Neurosci 1993;5:1413–20.

[16] d'Hooge R, Nagels G, Franck F, Bakker CE, Reyniers E, Storm K, et al. Mildly impaired water maze performance in male Fmr1 knockout mice. Neuroscience 1997;76:367–76.

[17] Dalm S, Grootendorst J, de Kloet ER, Oitzl MS. Quantification of swim patterns in the Morris water maze. Behav Res Methods Instrum Comput 2000;32:134–9.

[18] Deutsch JA. Spatial-learning in mutant mice. Science 1993;262:760–1.

[19] Francis DD, Zaharia MD, Shanks N, Anisman H. Stress induced disturbances in Morris watermaze performance: interstrain variability. Physiol Behav 1995;58:57–65.

[20] Freund JE. Mathematical statistics. 5th ed. Englewood Cliffs, NJ: Prentice Hall; 1992.

[21] Frick KM, Stillner ET, Berger-Sweeney J. Mice are not little rats: species differences in a one-day water maze task. Neuroreport 2000;11:3461–5.

[22] Garcia MF, Gordon MN, Hutton M, Lewis J, McGowan E, Dickey CA, et al. The retinal degeneration (rd) gene seriously impairs spatial cognitive performance in normal and Alzheimer's transgenic mice. Neuroreport 2004;15:73–7.

[23] Gerlai R. Ethological approaches in behavioral neurogenetic research. In: Crusio WE, Gerlai R, editors. Handbook of molecular-genetic techniques for brain and behavior research. Amsterdam: Elsevier; 1999. p. 605–13.

[24] Goldowitz D, Koch J. Performance of normal and neurological mutant mice on radial arm maze and active avoidance tasks. Behav Neural Biol 1986;46:216–26.

[25] Grant SGN, Odell TJ, Karl KA, Stein PL, Soriano P, Kandel ER. Impaired long-term potentiation, spatial-learning, and hippocampal development in Fyn mutant mice. Science 1992;258:1903–10.

[26] Grasmuck V, Desor D. Behavioural differentiation of rats confronted to a complex diving-for-food situation. Behav Processes 2002;58:67–77.

[27] Grubb SC, Churchill GA, Bogue MA. A collaborative database of inbred mouse strain characteristics. Bioinformatics 2004;20:2857–9.

[28] Guastavino J-M. Environmental features determining successful rearing in the mutant mouse *staggerer*. Physiol Behav 1984;32:225–8.

[29] Iida R, Yamada K, Mamiya T, Saito K, Seishima M, Nabeshima T. Characterization of learning and memory deficits in C57BL/6 mice infected with LP-BM5, a murine model of AIDS. J Neuroimmunol 1999;95:65–72.

[30] Iivonen H, Nurminen L, Harri M, Tanila H, Puoliv¨ali J. Hypothermia in mice tested in Morris water maze. Behav Brain Res 2003;141:207–13.

[31] Jones BJ, Roberts DJ. Quantitative measurement of motor incoordination in naive mice using an accelerating rotarod. J Pharm Pharmacol 1968;20:302–4.

[32] Kempermann G, Gage FH. Genetic determinants of adult hippocampal neurogenesis correlate with acquisition, but not probe trial performance, in the water maze task. Eur J Neurosci 2002;16:129–36.

[33] Klapdor K, VanderStaay FJ. The Morris water-escape task in mice: strain differences and effects of intra-maze contrast and brightness. Physiol Behav 1996;60:1247–54.

[34] Leighty RE, Nilsson LNG, Potter H, Costa DA, Low MA, Bales KR, et al. Use of multimetric statistical analysis to characterize and discriminate between the performance of four Alzheimer's transgenic mouse lines differing in Abeta deposition. Behav Brain Res 2004;153:107–21.

[35] Lewis RA. Predator mouse. Fly Anglers OnLine 2005:8.

[36] Logue SF, Paylor R, Wehner JM. Hippocampal lesions cause learning deficits in inbred mice in the Morris water maze and conditioned-fear task. Behav Neurosci 1997;111:104–13.

[37] Morris R. Developments of a water-maze procedure for studying spatial-learning in the rat. J Neurosci Methods 1984;11:47–60.

[38] Nakazawa K, Quirk MC, Chitwood RA, Watanabe M, Yeckel MF, Sun LD, et al. Requirement for hippocampal CA3 NMDA receptors in associative memory recall. Science 2002;297:211–8.

[39] Nowak RM, Paradiso JL. Walker's mammals of the world. 4th ed. Baltimore: Johns Hopkins University Press; 1983.

[40] Owen EH, Logue SF, Rasmussen DL, Wehner JM. Assessment of learning by the Morris water task and fear conditioning in inbred mouse
strains and F1 hybrids: implications of genetic background for single gene mutations and quantitative trait loci analyses. Neuroscience 1997;80:1087–99.

[41] Paigen K, Eppig JT. A Mouse Phenome Project. Mamm Genome 2000;11:715–7.

[42] Panda S, Provencio I, Tu DC, Pires SS, Rollag MD, Castrucci AM, et al. Melanopsin is required for non-image-forming photic responses in blind mice. Science 2003;301:525–7.

[43] Roder JK, Roder JC, Gerlai R. Memory and the effect of cold shock in the water maze in S100 beta transgenic mice. Physiol Behav 1996;60:611–5.

[44] Rustay NR, Wahlsten D, Crabbe JC. Assessment of genetic susceptibility to ethanol intoxication in mice. Proc Natl Acad Sci USA 2003;100:2917–22.

[45] Rustay NR, Wahlsten D, Crabbe JC. Influence of task parameters on rotarod performance and sensitivity to ethanol in mice. Behav Brain Res 2003;141:237–49.

[46] Semo M, Peirson S, Lupi D, Lucas RJ, Jeffery G, Foster RG. Melanopsin retinal ganglion cells and the maintenance of circadian and pupillary responses to light in aged rodless/coneless (rd/rdcl) mice. Eur J Neurosci 2003;17:1793–801.

[47] Stavenezer AJ, Hyde LA, Bimonte HA, Armstrong CM, Denenberg VH. Differential learning strategies in spatial and nonspatial versions of the Morris water maze in the C57BL6J inbred mouse strain. Behav Brain Res 2002;133:261–70.

[48] Steinberger D, Reynolds DS, Ferris P, Lincoln R, Datta S, Stanley J, et al. Genetic mapping of variation in spatial learning in the mouse. J Neurosci 2003;23:2426–33.

[49] Tarantino LM, Gould TJ, Druhan JP, Bucan M. Behavior and mutagenesis screens: the importance of baseline analysis of inbred strains. Mamm Genome 2000;11:555–64.

[50] Tecott LH, Nestler EJ. Neurobehavioral assessment in the information age. Nat Neurosci 2004;7:462–6.

[51] 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;58:213–8.

[52]   Upchurch M, Wehner JM. Differences between inbred strains of mice in Morris water maze performance. Behav Genet 1988;18:55–68.

[53]   van der Staay FJ. Effects of the size of the Morris water tank on spatial discrimination learning in the CFW1 mouse. Physiol Behav 2000;68:599–602.

[54]   V̄oikar V, Koks S, Vasar E, Rauvala H. Strain and gender differences in the behavior of mouse lines commonly used in transgenic studies. Physiol Behav 2001;72:271–81.

[55]   Wahlsten D. Sample size requirements for experiments on laboratory animals. In: Morm`ede P, Jones BC, editors. Neurobehavioral genetics: methods and applications. 2nd ed. Boca Raton, FL: CRC Press; 2005.

[56]   Wahlsten D, Crabbe JC, Dudek BC. Behavioral testing of standard inbred and 5HT1B knockout mice: implications of absent corpus callosum. Behav Brain Res 2001;125:23–32.

[57]   Wahlsten D, Metten P, Crabbe JC. A rating scale for wildness and ease of handling laboratory mice: results for 21 inbred strains tested in two laboratories. Genes Brain Behav 2003;2:71–9.

[58]   Wahlsten D, Metten P, Phillips TJ, Boehm II SL, Burkhart-Kasch S, Dorow J, et al. Different data from different labs: lessons from studies of gene–environment interaction. J Neurobiol 2003;54:283–311.

[59]   Wahlsten D, Metten P, Crabbe JC. Survey of 21 inbred mouse strains in two laboratories reveals that BTBR T/+ *tf/tf* has severely reduced hippocampal commissure and absent corpus callosum. Brain Res 2003;971:47–54.

[60]   Wahlsten D, Rustay NR, Metten P, Crabbe JC. In search of a better mouse test. Trends Neurosci 2003;26:132–6.

[61]   Whishaw IQ, Tomie J-A. Of mice and mazes: similarities between mice and rats on dry land but not water mazes. Physiol Behav 1996;60:1191–7.

[62]   Williams MT, Morford LL, McCrea AE, Wood SL, Vorhees CV. Administration of D, L-fenfluramine to rats produces learning deficits
in the Cincinnati water maze but not the Morris water maze: relationship to adrenal cortical output. Neurotoxicol Teratol 2002;24:783– 96.

[63]   Wolfer DP, Madani R, Valenti P, Li H-P. Extended analysis of path data from mutant mice using the public domain software Wintrack. Physiol Behav 2001;73:745–53.

[64]   Wolfer DP, M̈uller U, Staglier M, Li H-P. Assessing the effects of the 129/Sv genetic background on swimming navigation learning in trans-
genic mutants: a study using mice with a modified [i-amyloid precursor protein gene. Brain Res 1997;771:1–13.

[65] Wolfer DP, Stagljar-Bozicevic M, Errington M, Li H-P. Spatial memory and learning in transgenic mice: fact or artifact? News Physiol Sci 1998;13:118–22.

[66] Yoshida M, Goto K, Watanabe S. Task-dependent strain difference of spatial learning in C57BL/6N and BALB/c mice. Physiol Behav 2001;73:37–42.