Chapter 10. Neurobehavioral Testing

Specifications for the Conduct of Toxicity Studies by the Division of Translational Toxicology at the National Institute of Environmental Health Sciences

From: Roberts GK, Stout MD, editors. Specifications for the Conduct of Toxicity Studies by the Division of Translational Toxicology at the National Institute of Environmental Health Sciences. Research Triangle Park, NC: National Institute of Environmental Health Sciences; 2023. https://doi.org/10.22427/NIEHS-00

10. Neurobehavioral Testing

M.V. Behl[1], G.J. Harry[1], G.K. Roberts[1], M.D. Stout[1], S.K. Witchey[1] (Editors)

[1]Division of Translational Toxicology, National Institute of Environmental Health Sciences, Research Triangle Park, North Carolina, USA

Last Updated: March 2023

10.1. General Guidelines for Conducting Neurobehavioral Studies

Behavioral testing is a critical tool in assessing nervous system disorders. There are many well-established tests for evaluating various behavioral test results; however, details of how these tests are conducted and how environmental and experimental design factors are controlled are critical to providing reproducible data (Graham et al. 2018; Mandillo et al. 2008; Wahlsten 2011).

It is also known that different outcomes can arise from laboratory conditions extraneous to the test protocol (Crabbe et al. 1999; Harry et al. 2022; Kafkafi et al. 2005; Kafkafi et al. 2018; Lewejohann et al. 2006; Mandillo et al. 2008). These factors have been discussed in various publications and provide the basis for many of the details of the current guidelines (Genzel 2021; Hånell and Marklund 2014; Saré et al. 2021). To address the issue of test quality and consistency, a number of guidance documents, protocol publications, and data analysis papers have been published over the years (Bailey et al. 2006; Hånell and Marklund 2014; IPCS 1986; Moser 2011; NAFTA-TWG 2016; Saré et al. 2021; Slikker et al. 2005; Vorhees and Williams 2021).

Specific aspects of the handling of rodents during mating, pregnancy, and lactation are provided in Chapter 6 of the Specifications (Laboratory Animal Medicine and Toxicology). Specific timing for clinical observation assessments will be provided in a protocol outline directly to the contractor. Neuropathology requirements are described in Chapter 8 of the Specifications (Anatomic Pathology).

10.1.1. Housing

- During periods of gestation and lactation, crinkled Kraft paper shall be used in cages of rat dams. At each cage change occurring between time of birth to time of weaning, new cage bedding will be enhanced with a sampling (approximately one-fourth cup) of the original litter bedding to provide home-cage olfactory cues to minimize stress associated with cage change. This practice is specific to neurobehavioral studies and differs from standard Laboratory Animal Medicine and Toxicology practices of removing enrichment between GD 19 and PND 4 (Section 6.2.13, Environmental Enrichment).
- All animals at the age of weaning and older shall be group housed (by sex and dose group) per cage size guidelines and study needs. Male and female animals shall be housed similarly. Individual housing of mice shall be considered in the presence of aggressive behavior. If an animal is single housed, new cage bedding shall be enhanced with a sampling (approximately one-fourth cup) of the original litter, and a form of cage enrichment shall be used to minimize the stress of isolation. Scientific

justification for deviation from group housing shall be provided by the Contracting Officer's Representative (COR).

10.1.2. Animal Identification and Selection

- If a study design requires individual identification of preweanling pups, a temporary marking (e.g., paw tattoo, tail marking) shall be implemented in a manner to not cause damage that can compromise behavior (e.g., toe/paw damage that may compromise strength or motor-dependent behaviors). Temporary markings shall be replaced by a more permanent marking system upon weaning. If damage occurs (as confirmed by veterinary staff), the animal should be excluded from behavioral assessments due to possible effect on performance.
- For developmental exposure, litters shall be standardized on postnatal day (PND) 4 to 8–10 pups with specific sex distribution determined by the study protocol.
- Unless specifically noted for identified endpoints, for gestational and/or lactational exposure, one animal/sex/litter shall be randomly selected for any specific cohort for a specific behavioral test or sequence of behavioral testing.
- One animal can undergo more than one type of assessment according to a testing schedule ensuring no confounding across tests (e.g., tests are spaced at adequate intervals [days]; tests involving aversive stimuli [e.g., shock] shall be conducted at the end of the testing sequence).
- As determined by the study protocol, more than one cohort of animals can be used for behavioral testing if the testing history is consistent within a cohort (e.g., all animals assessed for any one endpoint will have had identical testing history in that all will have experienced the same test sequence).
- If, for any reason, an animal is removed from study due to health issues, it may be replaced with a matched animal from that same litter and dose group, per agreement of the COR.
- A process shall be put in place to ensure experimenter blinding of animal exposure status while running studies.

10.1.3. Randomization and Counterbalancing

- Animals shall be randomly selected for assignment to the behavioral testing cohort(s).
- All testing shall be counterbalanced for dose across testing apparatus and order of testing. If the same animal is to be tested multiple times for an activity, it shall be placed within the same apparatus for each test session.
- Males and females shall be tested at separate times within any single day. If the study design requires counterbalancing for sex, that information will be provided by the COR.
- If it is necessary for multiple technicians to perform any one task, the assignment of technicians shall be counterbalanced across dose groups to ensure an equal distribution of dose groups across technicians. For any observational endpoints, this requires technicians with a priori demonstrated >80% interrater reliability as

statistically determined from the ratings of two or more technicians on a specific test. For endpoints that can be significantly influenced by the technician (e.g., grip strength, rotarod), consistency of >80% inter-experimenter reliability across the actual technicians to conduct the study shall be demonstrated prior to the study (≤1 year).

- If an animal is terminated prior to testing, a secondary animal may be identified as a replacement and noted as such. The decision shall be made in consultation with the COR.

10.1.4. Animal Clinical Observations

- Body weights shall be recorded as specified in the study protocol.
- Formal (out-of-cage) clinical observations (characterized as "routine" or "clinical observations" in NTP Provantis) for clinical signs of toxicity will be recorded as specified in the study protocol. Clinical observations considered related to chemical exposure and observed at times other than scheduled observations will also be recorded in NTP Provantis.
- Additional nonexposure-related observations, outside of clinical observations, which may affect behavioral testing, shall be recorded in NTP Provantis. Examples include, but are not limited to, loss of toenail, injury to limb and/or tail, dropped to floor, or seizure during testing. Observations should be recorded in such a manner as to easily reflect testing days during review. Animals may continue in behavioral testing if their injury does not affect their overall well-being. The COR shall be notified of such conditions as they may influence study outcomes.

10.1.5. Selection of Testing Equipment and Procedures

- Commercially available equipment that is based on well-established methods for assessing the various neurobehavioral endpoints as demonstrated by published literature shall be used.
- Documentation shall be provided indicating that each test, as conducted by the testing facility (animal handling, environment, testing procedures and conditions), generates data reflecting a normal expected (per published literature) pattern of behavior of naive age-, sex-, and species-specific animals relevant to the study. This documentation can be provided within the form of historical control data captured on the equipment within the last 5 years. If data are not available, then prior to initiating testing of any study animals, a pilot study shall be run to confirm assay setup and laboratory proficiency. A cohort of five naive males (and five females, if females are to be assessed in the study) shall undergo each behavioral test at the defined ages.
- If there is a requirement to test for an acute (not developmental) effect on a specific behavior, a study using a positive control may be required to be performed prior to approval of the test paradigm. The specific experiments required prior to study initiation will be at the direction of the COR.
- Expected results shall have been demonstrated in control animals (species, strain, and sex) within approximately 1 year from the start of the study.

- Historical data, basic response, and positive control data for the most recent 5 years shall be available to the COR.
- All equipment and testing paradigms shall be approved by the COR prior to the start of the study.
- Detailed physical descriptions of the test equipment shall be reported (and can be provided via the commercial source manual). Details shall be supplied from each commercial supplier regarding definition of terms, recommended calculations of data (e.g., ambulatory activity), and calibration. Methods of calibration of each testing apparatus and identification of when and how often calibration is conducted shall be provided.

10.1.6. Handling for Specific Behavioral Tests

- Before behavioral testing, animals shall be acclimated with handling to ensure they do not undergo undue stress. Proper handling of rats consists of scooping up the animal with one hand under the chest and supporting the bottom with the other hand (UNC Basic Rat Handling and Technique Guide[1]). Mouse handling is improved by either nonaversive tunnel or cupping methods (Gouveia and Hurst 2017; Marcotte et al. 2021; Sensini et al. 2020). If tail handling is required, it must occur at the base of the tail. Tail handling is stressful to animals, so it should be used minimally.
- Further acclimation to any specific type of handling required for a behavioral test (e.g., grip strength, startle restraint) shall occur prior to testing to minimize handling stress at time of testing. This handling can be accomplished by acclimation to unique handling required for placing the animal in the restrainer used in the startle apparatus or on strain gauges for grip strength analysis.
- Consistency shall be maintained across animals within any specific test in terms of technicians' handling of the animal, placing of the animal within a test apparatus, removal of the animal from an apparatus, and returning the animal to its home cage.

10.1.7. Minimizing Olfactory Stimuli

- To minimize the influence of olfactory cues, the test environment shall be devoid of all specific odors to the extent possible, including, but not limited to, odors from chamber cleaning solutions and from the experimenter (e.g., perfumes, tobacco smoke, hand sanitizer, food).
- Animal-specific odors as they relate to stress, urine trail, or sex require minimization. Between animals (of the same sex) test arenas or restraints shall be wiped clean using a mild fragrant-free detergent with a disinfectant (e.g., Nolvasan [chlorhexidine diacetate] followed by a distilled water rinse). At completion of all testing for the day, any portion of the test apparatus that was in contact with the animal shall be cleaned with an excess of a mild fragrant-free detergent with a disinfectant, allowing the solution to remain on the apparatus for approximately 1 minute before wiping with distilled water, and wiped dry. For the Morris water maze (MWM), disturbance of the water between animals will minimize any urine trail.

[1] https://research.unc.edu/wp-content/uploads/sites/61/2020/12/rat-handling-and-techniques.pdf

- The influence of sex-specific odors (e.g., urine) on behavior shall be minimized. Unless otherwise requested that the experimental design include full counterbalancing for sex, testing each sex on either separate days or separate times of day with enhanced cleaning between sexes can minimize the influence of sex-related odors. Before handling animals of different sexes for behavioral testing, the experimenter shall change gloves and change or wipe down protective body coveralls to remove any traces of influencing odors.

10.1.8. Behavioral Testing Environment

- Outside of the home-cage room, lack of access to drinking water and food during the daytime shall not be >6 hours. Maintaining dosing via food or drinking water over the period outside of the home cage shall be determined by the study protocol.
- Ambient noise level shall be minimized in the testing facility. Testing rooms shall be maintained with white noise. A random examination shall confirm that the white noise (62–70 dB) is consistent across test units closest to and furthest from the white-noise speaker. Any occurrence of a loud noise during testing shall be recorded.
- If animals are tested in a sound-attenuating chamber with an individual background noise generator (e.g., startle apparatus) and the testing room is within a quiet area of the laboratory, considerations of not having full white noise for the room will be made by the COR.
- Conversation between experimenters within testing rooms shall be kept at a minimum.
- Quiet shall be maintained when placing animals into a test apparatus and during a test session.
- The testing units shall be stabilized to minimize any vibration, as applicable. Assessment of vibration of the testing room due to mechanical activity shall be conducted and confirmed absent or within a specific time of day/week for adjustment.
- Only the test cohort undergoing assessment shall be in the test room at any one time (e.g., animals not within the startle apparatus during a test session shall not be within the test room; animals not being tested for motor activity on that specific day shall not be within the test room). An exemption can be made for long-term learning and memory tests, such as the MWM.

10.1.9. Time of Testing

- All scheduled behavioral testing shall be conducted at an interval within 1 hour of a standard 12-hour light/dark cycle. Thus, testing will not commence until 1 hour after lights are turned on and will end by 1 hour before lights are turned off. The time of day of testing for each animal will be recorded. This time restriction does not apply to transport of animals to/from the testing room.
- Any shorter time interval required due to study design requires approval by the COR.
- Time of testing shall be counterbalanced across exposure group (for each sex per study design). Any alternative specific order for each test shall require COR approval.

- Any endpoints of repeated measures shall be conducted at approximately the same time of day for each animal.
- The effects of a direct dosing of a compound on the behavioral performance shall be considered, and the schedule of dosing to testing shall be coordinated to minimize confounding of such effects.
- Determination of the time of testing relative to the time of last dose will be made by the COR.
- When possible, behavioral testing shall be conducted before any direct dosing (e.g., gavage, injection, dermal application, inhalation) for that day unless study design is targeted to examine acute effects of chemical delivery or examine the peak time of effect.
- The timing of cage changes shall be controlled to ensure animals have a minimum of 24 hours post-cage change prior to testing.

10.1.10. Transport and Location of Animals for Pretest Holding

- Animals shall be transferred from their home-cage room and placed in a quiet holding area close to the testing room and outside of any high-traffic areas. Before the first behavioral test (open field), the animals will have become familiar with cage rack movement (down the hall or in the home room).
- Animals shall be transported and held in the holding area for a minimum of 30 minutes prior to initiating the first test of the day and for a minimum of 10 minutes if sequentially transported to a separate room for testing. Transport of animals is not restricted by the 1-hour lights-on/lights-off time interval required for behavioral assessments.
- Animals shall not be held within the testing room for tests that include a stimulus (e.g., startle), as nonspecific exposure would occur to animals outside of the test chamber.
- Animals shall not be held within the testing room for tests in which light level is a factor (i.e., open-field activity).
- Animals can be held within the testing room for the MWM, per conditions stated in the MWM description.

10.1.11. Retrieval of Animals after Test

- Animals shall remain in the test apparatus (e.g., activity chamber, startle apparatus, rotarod) until the test session times out for all animals to minimize disturbance and distraction.
- For the MWM, animals shall be gently wiped off upon removal from the water tank. Care shall be taken to not subject wet animals to air drafts and to minimize the potential for hypothermia.
- Animals shall be removed from the test apparatus and placed in the transport cage prior to returning to the home cage. If animals are transported in the home cage, they can be returned to that cage (without any requirement of a specific transport cage).

- The “tested” animals shall not be returned to a cage containing animals that have not been tested but shall be placed in a holding cage until all animals in the cage have been tested for the day.

10.1.12. Order of Behavioral Assessments

The initial assessments will acclimate the animal to handling, and assessments in the open field for motor activity will allow the animal to experience a novel environment and allow for detection of any severe motor deficits that may compromise subsequent behavioral assessments. Grip strength and rotarod training assess motor strength and coordination as well as learning. They are noninvasive and can be conducted in sequence prior to or after the startle response, depending on study design. The startle response is primarily a reflex response; however, there are different types of habituation (intrasession and across sessions) that require consideration if more than one testing session is conducted. Different tests for learning and memory can be integrated at different ages and is often the final test of the sequence. The inclusion of additional tests will require consideration of testing history and possible interference across tests.

- In general, excluding clinical observations, only one specific test will occur in any single day.
- Tests should be performed within a short time window to minimize variability associated with age or time from dosing.
- Behavioral assessments shall be performed in a sequential manner that shall not interfere with performance on subsequent tests (e.g., any test involving aversive stimuli, such as shock, shall be conducted at the end of the testing sequence or in a separate set of animals).
- Time intervals between tests shall be consistent across animals.

10.1.13. Age of Behavioral Assessments

Comprehensive evaluation for potential neurotoxicity includes assessments performed prior to weaning, and during the adolescent and adult life phases. A critical determination regarding the actual age of testing for a young animal is the time of weaning. All recommendations are to refrain from behavioral testing within 24 hours postweaning. The below age ranges for juvenile animals are based on the Division of Translational Toxicology practice of weaning at PND 28.

- Assessment of animals <60 days of age requires a very small window of age for testing to account and control for the developmental process of neural circuitry controlling such behavior. Assessment of animals as adults allows for a broader range in age but should be within 65–80 days of age, or 80–90 days of age and counterbalanced across groups.
- Assessment of preweaning behaviors, such as motor activity, requires age-appropriate testing apparatus to provide sufficient sampling of the behavior.
- Assessment of startle and prepulse startle inhibition (PPI) shall adhere to the developmental ontogeny of the associated neural circuitry and use equipment appropriate for the size and weight of the animal.

- Motor activity shall be measured in juveniles at PND 31 (±2) (rats)/PND 24 (±2) (mice) and adults PND 60 (±3) (rats/mice).
- Startle response and PPI shall be conducted at PND 32 (±3) (rats)/PND 25 (±3) (mice), and during adulthood PND 62 (±3) (rats/mice).
- MWM shall be conducted at PND 68 (±4) (rats/mice).
- *Additional behavioral testing, if required by the study protocol:*
    - Forelimb/hindlimb (FL/HL) grip strength shall be conducted in juveniles at PND 32 (±2) (rats)/PND 25 (±3) (mice) and/or adults at PND 60 (±3) (rats/mice).
    - Rotarod shall be conducted in juveniles at PND 32 (±3) (rats)/PND 25 (±3) (mice) and/or adults at PND 60 (±3) (rats/mice).
- It is recommended that the additional observational endpoints listed be included in the standard observational assessments taken over the course of the study.
    - Note: The above time frames are assuming animals are weaned at PND 28 (rats)/PND 21 (mice); relative adjustments shall be made if weaning occurs at a different time point.

10.2. Data Collection and Transmittal

- Data sheets shall include individual animal identifiers: litter (dam) number, sire number, pup number, dose group, sex, age, test date, apparatus identifier, and experimenter identifier. Dose group shall be coded in a manner to maintain experimenter blinding.
- For all computer-assisted tests, software-generated files describing the configuration of the test will be provided (these will include all parameters and units of measures and any time intervals and intensities used).
- For all computer-assisted tests, individual animal raw data files of all endpoints shall be submitted. These include arena maps, endpoints, and pathway tracking data.
- Video-capture images and/or pathway tracking and associated software files shall be provided in a format accessible by NIEHS.
- For each endpoint, all original raw data shall be provided in addition to the Excel or CSV files of specific endpoints.

10.3. Data and Statistical Requirements

- All information related to data processing, analysis, and outputs described in this section shall be provided in an electronic format; however, the specific disposition can be discussed with the COR on a study-specific basis. Examples of specific files include, but are not limited to, outlier analyses and justification for removed data, as well as graphical approaches for assessing normality (e.g., model residual plots, boxplots).
- Procedures for outlier identification/removal should be clearly described.
- The final statistical procedures used for all analyses should be clearly described.

10.3.1. One Measurement per Animal

- Account for censoring using time-to-event/survival models such as Cox proportional hazards modeling. Time-to-event modeling should be used for the rotarod and MWM analyses. (Do not remove censored values from the analysis.)
- Account for deviations from normality and heteroscedasticity across groups using the proper modeling approach (nonparametric Kruskal Wallis analysis of variance [ANOVA], data transformation and parametric ANOVA, or generalized linear modeling).
    - If the data are approximately normally distributed and the variances are similar between different groups, parametric ANOVA is recommended.
    - If parametric assumptions are not reasonable, use a nonparametric approach. Since the nonparametric Kruskal Wallis ANOVA assumes that the distributions of the different groups have a similar form, this assumption should be investigated by plotting the raw data distributions.
    - For count data, a Kruskal Wallis test may be sufficient, but a generalized linear model may be needed if the distributions from different groups are very different from one another. See below for more details on assessing distributional assumptions.
- Use the Shapiro-Wilk (SW) test to test for departures from normality and use Levene's test to test for heteroscedasticity. Graphical approaches using raw data should also be used to help inform whether to use a nonparametric statistical analysis approach. There is no need to rely solely on the SW test or Levene's test to assess parametric assumptions. Expert judgment may be used about the distributional assumptions based on the plots and the SW test and Levene's test results (Holson et al. 2008). Some examples related to determination of normality:
    - Graphical methods should be used to check for substantial deviations from normality and heteroscedasticity (e.g., model residual plots, normal quantile-quantile plots, and boxplots) because the SW test and Levene's test can be restrictively conservative or liberal for determining suitability of the data for ANOVA.
    - It is not necessary to account for unequal variance between groups in modeling if variance in the group with the smallest variance and variance in the group with the largest variance is within a factor of four of each other. A fourfold difference between variances is a general rule, not an absolute criterion, and should be considered part of expert judgment (Moore and McCabe 2001).
    - If there are very large differences in variance between treatment groups and the distributions seem to be very different from one another based on visual plots, then statistical methods should also account for heteroscedasticity as well as departures from normality (e.g., generalized linear model).
- Display data as mean/standard error of the mean (SEM), as well as plots of raw data with boxplots. The boxplots should help determine whether the distributions have a

similar form, find large deviations from normality (e.g., censoring or heavy skew), or indicate whether there is very large heteroscedasticity.

10.3.2. Multiple Measurements per Animal

- Similar to the above, account for censoring by using time-to-event modeling.
- All study designs should be appropriately counterbalanced during conduct, which should be either stated or described in the statistical analysis methods summary.
- Account for deviations from normality or heteroscedasticity. If the data distributions are very different from normal or if there is a large amount of heteroscedasticity, log transformation of the data should be considered before applying repeated measures ANOVA (RMANOVA). Assuming the study design is appropriately counterbalanced, the nonparametric Friedman test for continuous data can be used for nonnormal data. If the Friedman test is not appropriate, then generalized linear modeling can be used for count data.
- Use the SW test, Levene's test, ***and*** graphical methods to look at data distributions and help make decisions about whether to use RMANOVA or nonparametric Friedman (or another approach). It is acceptable to emphasize the plots for assessing the distributional assumptions and not rely solely on the SW test or Levene's test results.
- Display data as mean/SEM ***and*** boxplots to help determine whether the distributions are dissimilar between groups, are not normal, or have substantial heteroscedasticity. These plots can be placed in an appendix or delivered using another electronic format (see above).

10.3.3. Analyses for Sex Effects

- The study design must be appropriately counterbalanced during conduct. Counterbalance time of testing across dose levels and sex to avoid confounding. Counterbalancing is needed to include separate factors in statistical modeling.
- Present the data separately for each sex in data plots. To test for sex effects, use the two approaches described below. Each question compares the larger model to the smaller model using the likelihood ratio test. For each likelihood ratio test result (for each question), the value of the test statistic, the degrees of freedom for the test, and the p value should be presented. No additional statistical analysis or investigation is needed, even if the full model is shown to be more appropriate.

Question 1

Effect of SEX: Compare the Full to Simple Model (get a p value for likelihood ratio test).

Full Model: R = dose + time + sex + dose:time + time:sex + dose:sex + dose:time:sex + ~litter

Simple Model: R = dose + time + dose:time + ~litter

Question 2

Effect of INTERACTIONS: Compare the Full to Simple Model (get a p value for likelihood ratio test).

Full Model: R = dose + time + sex + dose:time + time:sex + dose:sex + dose:time:sex + ~litter

Simple Model: R = dose + time + sex + ~litter

10.4. Observational Assessments

10.4.1. Cage-side Observations in Juvenile Pups

Rodent juvenile activity accounts for many critical features of nervous system integration, including motor function, strength, and social behavior (Vanderschuren et al. 2016). Rats begin to play toward the end of the third week of life, peaking during the fourth and fifth weeks and then decreasing with sexual maturity (Pellis and Pellis 1990). Play by juvenile rats is often considered a potential indictor of good animal welfare (Oliveira et al. 2010) and has been demonstrated to be associated with the development of sociocognitive skills (Baarendse et al. 2013; Stark and Pellis 2020). It has been used as a model for analyzing neurodevelopmental disorders (Burke et al. 2017; Pellis et al. 2022). Play behavior can be scored as it occurs naturally in a litter (Lampe et al. 2019); however, this limits the level of detail that can be obtained. Yet, if allowed multiple opportunities for observation, home-cage assessments can effectively detect the presence/absence of play behavior but not the subtle differences in such behavior.

10.4.2. Scoring Juvenile Social Behavior

During normal scheduled clinical observations, juvenile social behaviors will be evaluated in group-housed animals over a 1-week period from PND 35 to PND 42. Pinning and pouncing are considered the main indices of social play behavior in rats because they strongly co-vary with other playful social behaviors, such as following and wrestling (Panksepp and Beatty 1980; Pellis et al. 2022; VanRyzin et al. 2020).

Play behavior will be assessed using a rating scale depicting the absence (1) or presence (2) of the specific behavior occurring within the home cage:

- Pinning behavior (without associated vocalizations)
- Pinning behavior (with associated vocalizations)
- Pouncing behavior
- Aggressive behavior (e.g., biting)

10.5. Locomotor Activity

An assessment of motor function by locomotor activity provides an indication not only of the activity level of the animal but, if decreased, an indication of concern for either general health or motor strength that could compromise subsequent behavioral evaluations. As an assessment of motor function, locomotor activity captured by automated photocell or video-capture detection systems allows for the evaluation of general motor activity; in addition, by using the stimulus of novelty, free exploration in the arena can be used to examine curiosity and exploration (Pisula and Modlinska 2020). Motor activity devices (photocell or video tracking) shall be capable of evaluating ambulatory motor activity in a two-dimensional manner (x, y planes) and rearing (z plane) in a time-interval manner and will be appropriate for the age and size of the animal. The

system shall be capable of measuring thigmotaxis and regional preference within the arena and of providing information on the ambulatory activity path length.

10.5.1. Configuration

Spontaneous locomotor activity shall be assessed within a defined arena using a commercially available system with documentation of usage and demonstration of biological response from positive control agents. The configuration of the system shall allow for data collection of activity within the entire arena, immediately along the perimeter at the chamber wall (margin zone), and within a defined smaller center area (center zone).

- Photocell open-field arena apparatus example: approximately 40 cm × 40 cm × 20 cm photocell device using a two-dimensional sensor array configuration (photocells at 1-inch intervals). Zones: margin (outermost infrared beam on each sensor, which would represent 1-inch margin at the wall of the arena) and center (5 × 5 inch) square. Video-capture systems would configure a similar arena and arena map for data capture. In addition, contrast distinction and lighting will be sufficient for uniform camera detection (e.g., white animal on a dark background).
- For photocell devices, the height placement of the photocell banks for horizontal measures shall be set to detect the midpoint of the body trunk as appropriate for the age and species of animal under study.
- For rearing behavior, the detection limit (photocell height, video recording height) will be at a height equivalent to at least three-quarters of the full rearing height of the animal to ensure accurate detection of full hindlimb rearing, exclusive of raising of the head, back, or slight rear, not requiring full weight to be placed on hindlimbs. (This height location shall be empirically determined for each species, strain, age, sex, and size of animal under study by confirmation of experimenter-observer counts compared with photocell detection).
- Parameters of the testing apparatus (height location of the photocells for ambulatory activity, empirically determined height location of photocell bank for rearing, arena zone definitions) shall be documented in the study file.
- At the beginning of each test session, each apparatus shall be calibrated for photocell alignment and function as instructed in the manufacturer's manual. In addition, each unit shall be tested to ensure accurate tracking by moving either the experimenter's hand or a controlled moving item (e.g., plastic ball) within the chamber and following the tracking pathway on the computer screen. The instrument shall provide a software diagnostic feature to be run prior to each session.
- Tests of motor activity shall be conducted under lighting conditions that maintain the normal home-cage room light/dark cycle. The lighting level within the front section of the home cage will be measured using a Lux meter. Standardization of lighting for each motor activity arena shall be similar to the level measured for the home cage. In configuration of the test room and apparatus placement, the uniformity of luminance shall be confirmed. The luminance shall be similar across each full arena to ensure no area of the arena is within shadows. Meeting these specifications may require modification of the actual luminance of the room lighting. The actual luminance of the room and in each of the arenas shall be tested using a Lux meter and recorded.

10.5.2. Initiation and Duration of Testing

- Animals shall be placed in the center of the activity arena. Each apparatus shall be programmed for the software to automatically start the test session with the detection of first movement within the chamber. If this is not possible, then the computer test session shall be started manually by the experimenter upon entry of animal into the chamber and clearance of the experimenter's hand.
- Standard Test: Measurements shall be collected in 5-minute epochs for a total of 30 minutes (preweanling/weanling <35 days of age) or 45 minutes (>45 days of age).

Endpoints for Collection

- For the entire arena, data shall be collected in 5-minute epochs for total activity, ambulatory activity (as defined by manufacture), fine movements, distance traveled, stationary time, rearing events, and time spent rearing.
- For the full arena, total session data shall be calculated for total activity, ambulatory activity, distance traveled, rearing events, and time spent rearing.
- In defined zones, data shall be collected in 5-minute epochs for ambulatory activity, distance traveled, time spent within defined zones (margin time [thigmotaxis]; center arena), and entries into zones.
- For the full arena, a pathway track or heat map for each animal shall be recorded for the entire session.
- Commercial supplier information defining how each endpoint is captured/determined and description of any calculations necessary for endpoint determination shall be provided in the study file.

10.5.3. Endpoints for Analysis

Total Session

- Full Arena: Total activity, ambulatory activity (as defined by manufacture), distance traveled, rearing events, and time spent rearing.
- Zones: Ambulatory activity, distance traveled, time in zone, and entries into zone.

Epochs

- Full Arena: Total activity, ambulatory activity, distance traveled, rearing events, and time spent rearing.
- Zones: Ambulatory activity.

10.5.4. Statistical Analysis

- Data will be examined for homogeneity of variance.
- Data transformations, such as logarithms, shall be considered for analysis over epochs, if necessary, to meet model assumptions.
- Data obtained across the full session, either in the full arena or in epochs, shall be analyzed with ANOVA, with dose as a factor.

- Data collected in 5-minute epochs shall be analyzed with RMANOVA, with dose and time as factors.

10.6. Forelimb and Hindlimb Grip Strength

To detect alterations within the peripheral nervous system or the spinal cord that would compromise limb strength or motor behavior, grip strength of the forelimbs and hindlimbs shall be assessed using a digital force gauge (Maurissen et al. 2003; Meyer et al. 1979; Takeshita et al. 2017).[2]

10.6.1. Configuration of Grip Strength Apparatus

- Animals shall be assessed for fore- and hindlimb grip strength using a strain-gauge system appropriate for the species and age of animal (screen, bar, overall length of platform, gauge strength size [kg]). Preference is given to equipment configuration that allows assessment of fore- and hindlimb grip strength in one pass and provides physical support for the animal.
- The apparatus shall be placed on a stable surface, away from drafts or vents that could disturb the measurement by the sensor.
- A bar shall be used for assessing fore- and hindlimb strength in adult rats. It is recommended that a screen grid not be used with adult rats due to the nature of the procedure and to prevent the error of the animal not being able to rapidly release the grid, thus altering the grip score.
- For juvenile rats and all mice, a screen grid or T-bar, sized appropriately, shall be used to assess fore- and hindlimb strength.
- Control animals should measure within the midrange of the meter, and the settings shall allow for detection of an increase or decrease in strength.

10.6.2. Grip Strength Protocol

- Confirm that the forelimb gauge is set to PULL mode and the hindlimb gauge is set to PUSH mode and that both gauges have been reset to zero.
- Place the animal on the center platform of the apparatus, facing the forelimb gauge. Set the forepaws on the screen (juvenile rats and mice) or bar (adult rats), attached to a strain gauge. Alternately, hold the animal by the "neck scruff" or base of the tail so only its front paws grip the grid platform/bar (handling method to be standardized across all animals in the test).
- Contact with the forelimb apparatus shall require all four digits of both limbs.
- Once a successful grip is observed, hold the animal by the base of the tail and gently pull horizontally and quickly with an even force in one continuous motion until its grip is released down the complete length of the grid/bar. The propensity is that the animal will cling onto the grid/bar until it can no longer resist the increasing force before it is released. Immediately following this and within the one smooth,

[2] https://www.mousephenotype.org/impress/ProcedureInfo?action=list&procID=1130; https://treat-nmd.org/wp-content/uploads/2016/08/MDX-DMD_M.2.2.001.pdf

continuous pulling motion for the forelimb assessment, the hindlimbs are allowed to grasp a bar (or grid) as the animal's body is quickly but smoothly pulled away. A slight pause may be necessary before beginning the pulling motion to assure the rat's digits are properly curled around the bar and the paws are not crossed.

- Do not allow the front feet to touch the hindlimb screen or bar.
- Do not allow the experimenter's hand to touch the screen or bar.
- Do not allow the toenails to catch in the mesh as this can result in a strong jerk motion and invalidate the reading.
- Three sequential trials shall be performed with an approximate minimum intertrial interval (ITI) of 10 seconds. Juvenile rats and mice may require a longer ITI to minimize fatigue, if more than three trials are required (to be determined by pilot study).
- The digital readouts on the gauge shall be recorded and the gauges reset.
- If necessary, two additional trials can be run to obtain three valid trials for averaging.
- If any animal fails more than two trials, that shall be noted.
- Unacceptable trials include: (a) the animal is aggressive and cannot be properly handled; (b) the animal fails to grip the mesh grid or T-bar with two paws; (c) the animal reacts in such a manner that the grid mesh or T-bar is released before the technician pulls the animal away from the gripping surface; (d) for some reason, the force gauge is accidentally activated by contact unrelated to the appropriate experimental procedure; and (e) the technician pulls the animal too strongly to mask animal response.

10.6.3. Endpoints for Collection

- Each individual gauge reading for forelimb grip strength
- Each individual gauge reading for hindlimb grip strength
- Average of three valid gauge readings for forelimb grip strength
- Average of three valid gauge readings for hindlimb grip strength
- If fewer than three valid readings, average of two valid trials can be calculated and the failed trials noted.

10.6.4. Endpoints for Analysis

- Mean forelimb grip strength for three valid trials
- Mean hindlimb grip strength for three valid trials

10.6.5. Statistical Analysis

- Data will be examined for homogeneity of variance.
- Data transformations shall be considered for analysis.
- Considerations of body weight difference as an influencing factor shall be made.

- Data obtained shall be analyzed with ANOVA, with dose as a factor.

10.7. Accelerating Rotarod (Motor Coordination)

The rotarod apparatus/protocol can be used to measure motor function, motor learning, coordination, and equilibrium in both rats and mice (Chapillon et al. 1998; Crawley 1999; Rustay et al. 2003a; 2003b). Assessment of motor coordination using the rotarod requires the animal to learn the novel task. To ensure all subjects have learned the task to the same degree, thus enabling experimenters to accurately measure differences in motor coordination and equilibrium, 2 weeks of training with three daily sessions of three trials per week are normally required in rats and mice.

10.7.1. Configuration of Rotarod Apparatus

- The automated apparatus shall be equipped with a rotating rod of a diameter appropriate for the size of animals tested. The rod will be grooved to allow for gripping by the animal.
- An apparatus that automatically records latency to fall and rotational speed (rpm) of the rod at time of fall under an accelerating or fixed speed is recommended.
- Before study initiation, data obtained within the previous year will be available to confirm the appropriate acceleration rate for the species, strain, sex, and age of the animal and the proficiency of the laboratory. The rate shall allow control animals to improve performance over trials without demonstrating a ceiling effect; the rate will not be so aggressive as to prevent control animals from adequately performing.
- Modifications to the apparatus can be employed to provide better traction/grip for the animal, with coarse rubber, Velcro, or fine grit (320 grit) sandpaper (this has been found to be helpful with mice or young rats) (Bohlen et al. 2009).
- The apparatus shall have partitions between animals and an enclosure to prevent animals from escaping after a safe landing.
- Calibration of rotational acceleration will be conducted within a year of study and data provided in the study report (Bohlen et al. 2009).

10.7.2. Rotarod Protocol

- Rats and mice may be placed on the rod by gripping the tail and/or by grasping the animal around the body. The limbs should be restrained as much as possible to minimize the animal grasping peripheral structures. An animal is placed on the rod, facing away from the direction of rotation, so it has to walk forward to stay upright. In mice, holding the animal by the tail at an angle of 40° below horizontal works well. On a multi-animal apparatus, animals are quickly placed on the rod, facing in the correct direction, and are in stable position at the start of timing.
- The first exposure to the rotarod serves as an initial training trial and is not included in the performance data analysis. With this first trial on the first day, the animal learns to balance on the stationary rod and then to maintain balance on the rotating rod. This first trial can be either (1) a rod constantly rotating at approximately 10 rpm for 1 minute or (2) on rotation parameters consistent with the testing trials. For the

training trial, the animals are placed on the low-speed rotating rod. If the animal falls off in less than a minute, it is immediately placed back on the rod until it stays on the rod for a full minute. The animal is allowed a 10-minute rest interval before the next trial.

- On day 1, the first exposure trial is followed by three trials. To ensure all subjects have learned the task to the same degree, thus enabling experimenters to accurately measure differences in motor coordination and equilibrium, 2 weeks of training with three daily sessions are normally required. Three trials shall be run per day, three times per week for 2 weeks, with a minimum ITI (rat: 10 minutes; mouse: 15 minutes) to minimize fatigue.
- ITI will be maintained across all animals. The rod will be rotating at 2 to 4 rpm (warm-up speeds). Any subject that falls or jumps from the rod during the first approximately 10 seconds of testing can be replaced, and the lane timer can be restarted.
- Overall test confounders include animals that cling to the rod but do not fall (passive rotations), animals that refuse the test and simply fall without any real evidence of altered muscle strength (exclude as outliers), and weight and size of animals—heavier and larger animals perform less well on the rotarod and fatigue with progressively longer latencies (confirm with fixed speed tests).

10.7.3. Endpoints for Collection

- Latency to fall
- Record the occurrence of jumping or passive rotations with speed or time of occurrence
- Record a latency of maximum cut-off for animals that did not fall off within the test interval

10.7.4. Endpoints for Analysis

- Latency to fall: individual trials and mean per day (excluding initial training trial)

10.7.5. Statistical Analysis

- The mean response of day 1 and day 2 shall be analyzed with ANOVA, with dose as a factor.
- Data across trials on each day shall be conducted with RMANOVA, with dose and trial as factors.
- The data across trials on each day shall be analyzed for pairwise comparison and trend across dose for each trial, using the F-test ANOVA.
- The mean response on day 1 and day 2 shall be analyzed by RMANOVA.
- If the data represent censored observations, the amount of censoring shall be reported and an appropriate analysis (e.g., time-to-event modeling) shall be considered.

10.8. Acoustic Startle Reactivity and Prepulse Startle Inhibition

The startle response is an unconditional reflex, characterized by the rapid contraction of skeletal muscles, in response to a sudden and intense startling stimulus (e.g., noise burst, air puff, light flash). In rodents, the acoustic startle response (ASR) can be used to study habituation, sensitization, classical conditioning, fear, and anxiety. Habituation to the startle response is a form of nonassociative learning and can also be viewed as a sensory filtering process as it decreases an organism's response to a nonthreatening stimulus. Habituation can be examined within a test session (short-term habituation) or across sessions (long-term habituation). Within a session, habituation normally occurs within the first 10 trials and over 4–5 days for across sessions (Pilz and Schnitzler 1996; Pilz et al. 2014; Valsamis and Schmid 2011). PPI describes the phenomenon in which a weak initial stimulus (prepulse) inhibits the startle response that is elicited by a strong stimulus. The level to which the prepulse stimuli inhibits the startle response increases with prepulse intensity. Animals shall undergo testing for startle response and PPI using a computer-assisted automated startle/PPI system. This system shall allow for the ability to view the continuous individual waveform responses and to examine the data post hoc based on shifting the millisecond time interval for recording a response. Recent articles on procedural methods and optimization considerations for ASR and PPI are available (Hormigo et al. 2019; Miller et al. 2021; Shoji and Miyakawa 2018; Valsamis and Schmid 2011).

10.8.1. Testing Units and Calibration

- All testing units shall be housed in individual sound-attenuated chambers within a testing room under normal animal facility environmental conditions.
- Calibration: Calibration of the sound (sound meter) and the movement (e.g., oscillation calibration device) sensors is critical for obtaining valid test results. A set background decibel level is essential when conducting experiments examining PPI as the prepulse levels are set relative to background. Thus, calibration shall be to a specific level and not to a range. This may require an extended time period to optimize calibration. It is recommended that the full calibration of each unit be scheduled to allow for adequate time prior to the start of any specific experiment. Once calibrated for offset (gain) and sound, the units should remain relatively consistent over time. If a unit is identified to drift over time, exclusion of that unit should be considered until recalibration from the commercial supplier. Confirmation of calibration shall be conducted within 24 hours prior to the start and 24 hours following cessation of a specific testing time for an experimental set of animals. Adjustments to sound or gain may be required based on species, strain, age, or animal tested.
    - Each sensing plate/unit shall be calibrated for mechanical and circuit offsets or gain using an oscillation calibration device following detailed instructions provided by the commercial supplier. The gain shall be set for each age/weight of the animal species and strain tested. Uniform readings (actual and not range) across units shall be confirmed. The actual value shall be recorded for each unit.
    - Using a sound meter, each unit shall be calibrated to the set background decibel level. The preferred level is 65 dB. Given the increase in fan noise

with aging equipment, if experimenters are unable to lower units to this level, then they will set the lowest level between 65 and 70 dB and record the background level. Set all units to the same background decibel. If, for any reason, a unit is unable to be set at the same background sound level, there can be a variability of ±1 dB for background across units (note how this will alter settings for prepulse decibels).

- The prepulse stimuli decibel levels will be set relative to the background decibel level for each unit. Using a sound meter, confirm the decibel level for each prepulse stimulus and for the 120-dB startle stimulus during the calibration stage. If calibration is required, balance the calibration for that unit across all decibel levels (background, prepulse, startle). Record any deviation.

10.8.2. Adaptation to Handling and Holding

- The startle apparatus requires that the animal be restricted for movement during the test session. The animal's species, age, and weight determine the appropriate size of the enclosure to restrict mobility as recommended by the manufacturer. The holding enclosure represents a novel environment for the animal and can alter an initial startle response. Unique handling is required to place the animal into the holder to minimize stress and allow for acclimation. For a tube type holder, the handling sequence involves gently squeezing the forepaws together so they cross on the underside of the animal, holding the hindquarters to prevent perambulation, introducing the animal into the holder or other suitable facsimile, and holding it there for 3 minutes (Geyer and Swerdlow 1998). In general, this procedure shall be conducted two times within the week prior to first startle test. For an open holding chamber, the animals can be placed into the holder and held for 3 minutes for one session within the week prior to the first startle.

10.8.3. Confirmation of Startle Decibel and Prepulse Intensities

- Before examination of animals on study, within 1 year of study initiation, the following will be conducted in animals (five/sex) for specific test ages, species, and strain to establish and confirm optimal intensities and absence of drift. Once established, these settings shall be confirmed on an annual basis prior to initiating testing on study animals.
    - An input/output function test shall be conducted. After a 5-minute acclimation period under constant background white noise of 65 dB, startle stimuli (20 milliseconds) shall be delivered on an ITI of 20 seconds. Startle stimuli events shall start at approximately 75 dB and increase by 5-dB increments until reaching 120–125 dB. Startle magnitudes shall be sampled each millisecond for 200 milliseconds beginning at the onset of the startle stimulus. These data shall be used to determine the maximum startle response (largest response within 200 milliseconds), provide information on the waveform, and average the response over the entire response window.
    - PPI stimulus intensities shall be identified to elicit intermediate levels of PPI to allow for treatment-induced increases or decreases in PPI to be observed. Maximum startle response to each of the individual PPI intensities shall be

determined. The required number and interval of PPI intensities shall be determined for the species, strain, and age of the animal.

10.8.4. Paradigm Configuration

- Background level: 65 dB
- PPI intensities set at 3, 6, 12, and 15 dB above background

10.8.5. Prepulse Auditory Startle Inhibition (PPI) Protocol

- Each session shall start with a 5-minute period of acclimation to the restrainer and chamber with continuous background noise (65 dB).
- Delivery of startle trials shall be under a fast rise time (<2 milliseconds) burst of noise presented for a 40-millisecond duration at an intensity of 120 dB.
- Trials will be delivered according to a variable ITI of 15 seconds, with a range of 7–23 seconds.
- The delivery of stimuli shall follow the sequence outlined in Table 10-1 that includes six initial 120-dB pulse-only trials followed by two "blocks" of trials that represent each of the paired prepulse intensities with a 120-dB trial five times each, five of the 120-dB pulse trials, and one to two no-stimulus trials presented in a pseudorandomized manner. These two blocks shall be followed by 10 of the 120-dB pulse trials.
- The full collection window shall be set at 500 milliseconds: (1) a sampling of 250 milliseconds preceding the 120-dB startle elicitation to confirm confounding with baseline activity and/or confirm absence of startle response elicited by the prepulse intensities followed by (2) a sampling of 100 milliseconds to collect the maximum startle response. Data for determining PPI will be obtained from the 100-millisecond sampling window measured from startle stimulus onset. The sampling during the 250 milliseconds preceding the 120-dB startle stimulus, but following the prepulse stimulus, shall be collected and examined to determine any group differences in response to each prepulse intensity.
- Prepulse stimuli (3, 6, 12, 15 dB above threshold unless empirically changed on the basis of the pilot study results) shall be presented for a 20-millisecond duration with an interstimulus interval of 65 milliseconds (mice) and 80 milliseconds (rats) before the onset of the 120-dB startle stimulus.
- Peak response magnitude (i.e., V_{max}) on no-stimulus (NOSTIM) trials shall be recorded and reported as a sampling of excessive activity of the animal within the chamber.

Table 10-1. PPI Protocol for Stimuli Delivery

Trial No.	Block No.	Trial Name	Prepulse Level (dB)
1	0	P120	–
2	1	P120	–
3	1	P120	–
4	1	P120	–
5	1	P120	–
6	1	P120	–
7	**2**	**NOSTIM**	–
8	2	PP77P120	77
9	2	PP68P120	68
10	2	PP80P120	80
11	2	P120	–
12	2	PP71P120	71
13	2	PP71P120	71
14	2	PP80P120	80
15	2	PP77P120	77
17	2	PP68P120	68
18	2	P120	–
19	2	P120	–
20	2	PP77P120	77
21	2	PP71P120	71
22	2	PP68P120	68
23	**2**	**NOSTIM**	–
24	2	PP80P120	80
25	2	PP80P120	80
26	2	P120	–
27	2	PP77P120	77
28	2	PP71P120	71
29	2	PP68P120	68
30	2	PP77P120	77
31	2	PP68P120	68
32	2	PP80P120	80
33	2	P120	–
35	2	PP71P120	71
36	3	P120	–
37	3	PP77P120	77
38	3	PP80P120	80

Trial No.	Block No.	Trial Name	Prepulse Level (dB)
39	3	PP68P120	68
40	3	PP71P120	71
41	3	P120	–
42	3	PP80P120	80
43	3	PP71P120	71
44	**3**	**NOSTIM**	–
45	3	PP77P120	77
46	3	PP68P120	68
47	3	PP77P120	77
48	3	PP71P120	71
49	3	PP80P120	80
50	3	PP68P120	68
51	3	P120	–
52	**3**	**NOSTIM**	–
53	3	PP71P120	71
54	3	PP80P120	80
55	3	P120	–
56	3	PP68P120	68
57	3	PP77P120	77
58	3	PP77P120	77
59	3	P120	–
60	3	PP68P120	68
61	3	PP71P120	71
62	3	PP80P120	80
63	4	P120	–
64	4	P120	–
65	4	P120	–
66	4	P120	–
67	4	P120	–
68	5	P120	–
69	5	P120	–
70	**5**	**NOSTIM**	–
71	5	P120	–
72	5	P120	–
73	5	P120	–

NOSTIM = no stimulus.

10.8.6. Endpoints for Collection

- For startle and prepulse startle inhibition, the schedule of the stimulus presentations shall adhere to Table 10-1, and documentation will be provided.
- Peak response amplitude (i.e., V_{max}) for each startle trial.
- Peak response amplitude for each prepulse stimulus intensity during the 250-millisecond sampling interval to determine whether there is an alteration in the prepulse threshold or a significant movement response that may affect the ASR.
- Time to maximum response (i.e., T_{max}) for each trial, excluding NOSTIM.
- If available: latency to onset of response, rise time of response as identified by the manufacturer's instructions.

10.8.7. Endpoints for Analysis

- First 120-dB trial: Peak response amplitude – representative of naive startle response.
- Peak response amplitude for individual animals and each 120-dB-only trial.
- The measured startle response is typically lognormally distributed across pulse types (Csomor et al. 2008) and often the median response over a block is more robust than the mean. Therefore, where applicable, both mean and median response will be calculated.
- Median peak response amplitude of 120 dB only occurring in the "blocks" of trials inclusive of prepulse stimulus intensities (Blocks 2 and 3) for each animal. This will be considered the "120-dB peak response" in calculating percent prepulse inhibition (% PPI) for that individual animal.
- Mean and median peak response amplitude for each prepulse intensity stimulus for each animal.
- PPI: For each individual animal and each PPI trial, the % PPI shall be calculated as [(120-dB peak response − prepulse peak response)/120-dB peak response] × 100. For each animal and each prepulse stimulus intensity, the average % PPI is calculated across all matched trials for the entire test session.
- Habituation: Calculated as the change in peak response amplitude over trials. Percentage change is calculated as a change of peak response of the last 120-dB trial compared with peak response of the first 120-dB trial for each individual animal.
- Data shall be represented as the actual metric recorded—e.g., Newtons (Amplitude [N]), static weights (Amplitude [g]), or volts (Amplitude [v])—per manufacturer's instructions.

10.8.8. Statistical Analysis

- Data shall be evaluated for assumptions of ANOVA (homogeneity of variance and normally distributed observations). Nonparametric methods should be considered if these assumptions are violated.
- Peak response amplitude of the first 120-dB ASR trial shall be analyzed using a one-way ANOVA.

- Peak response amplitude of 120 dB across trials shall be analyzed using RMANOVA (dose and trial as factors).
- Peak response amplitude for paired prepulse intensity per 120-dB trial for each prepulse intensity shall be analyzed by one-way ANOVA (dose as factor).
- Negative PPI values shall be set to 0 (the incidence of negative PPI values will be recorded).
- Calculated % PPI for each prepulse intensity shall be analyzed by one-way ANOVA (dose as factor).
- Calculated % habituation shall be analyzed with one-way ANOVA (dose as factor).

10.9. Morris Water Maze

Learning and memory will be assessed in the MWM using commercially available video tracking equipment and software that has demonstrated use in the published literature. Animals will undergo a sequence of training tests to evaluate performance in an MWM. On days 1 and 2, animals shall be familiarized to the tank, water, and swimming requirements of the test and assessed for nonspatial cued learning. This shall be followed by acquisition of a spatial hidden platform task (three training trials/day for 7 consecutive days). Twenty-four hours after completion of the hidden platform task, spatial reference memory shall be assessed in a probe trial. Forty-eight hours after completion of the probe trial, performance on a spatial reversal acquisition task shall be assessed (three trials/day for 3 days) (Gallagher et al. 1993; Gerlai 2001; Maei et al. 2009; Vorhees and Williams 2006).

10.9.1. Testing Environment

- Performance is dependent upon maze configuration. The use of multiple tanks requires uniformity of tank configuration and dimensional details of the room, a spatial-defined area, and spatial cues as they relate to the visual field of the animal. All specific details of physical properties of the tank configuration and cue placement shall be documented and provided in the study report. The tank shall be properly sized. The final report shall include all variables and detailed physical descriptions of the tank, platform, platform placement, and visual cues.

10.9.2. Tank and Platform

Tank

- The interior diameter of the standard circular tank shall be approximately 180 cm (6 ft.) (adult rats) or 130–150 cm (4–5 ft.) (immature rats and adult mice) with nonreflective interior surfaces. A large tank can be modified to the smaller size by the insertion of a circular ring to decrease diameter.
- When filled with water, the depth shall be sufficient to prevent the animal from touching the bottom (e.g., approximately 28–35 cm for a 70-day-old rat; 15–20 cm for mice) and with a surface-to-tank-lip distance of approximately 10–12 cm to prevent the animal from jumping out of the tank but allowing for line of sight to visual cues.

- The water temperature shall be equilibrated to the ambient room temperature (approximately 22°C). During training and testing, the water temperature shall be recorded at the start and end of each test day to confirm it is within this range.
- The tank shall have no prominent interior features that might provide proximal cues, such as markings inside the tank (e.g., welded seams shall be smoothed and painted to blend with the surrounding walls). The sides of the tank shall be smooth to minimize animals' attempts at climbing walls and should be a nonreflective matte finish. Location markers shall be placed on the exterior of the tank and will line up with markers on the floor to ensure identical location placement of the tank across trials/sessions.
- For albino (or light-coated) rodents, the tank shall be dark, and the platform shall be dark or clear to prevent visualization from the surface of the water. For pigmented rodents, a light-colored tank with a similarly colored or clear platform shall be used. If necessary, for animals with a dark coat, a white nontoxic water coloring (e.g., Crayola watercolor paint, tempera paint) can be used with a clear platform.
- The maze shall be designated into four equal quadrants (randomly identified as Northwest, Northeast, Southwest, and Southeast).

Platform

- The tank shall have the capability for using a hidden platform and a visible platform. The goal platform shall be approximately 10 cm in diameter for rats and 8 cm in diameter for mice. The platform shall be covered by a nonreflective textured material to allow for gripping by the animal to facilitate escape from the water.
- The platform shall be positioned in the respective quadrant (cued learning: NW; hidden platform: NE; reversal learning: SW), located approximately 36 cm (rat) or 30 cm (mice) from the interior wall of the tank (to minimize the chance of finding the platform simply by swimming along the tank wall), yet at a distance from the center. The platform shall remain within that quadrant for all animals during cued (visual) and hidden platform acquisition trials.
- Height of platform: For cued learning, the visible platform shall be at a height of approximately 1.5 cm above the surface of the water with a visible projection (10–12 cm^2 "flag") attached and rising approximately 13 cm above the visible platform. For spatial learning and reversal learning, the hidden platform shall be submerged below the surface of the water (approximately 1 cm for rats; approximately 0.5 cm for mice).
- The platform will be immobilized within the tank to prevent any shift over the day's testing.

Cleaning of Tank

- After each animal test, floating feces shall be removed and water dispersed to minimize urine scent near the platform or prior platform location.
- At a minimum, the tank shall be drained and rinsed clean at the end of each 5-day period, refilled, and allowed to equilibrate to ambient temperature. (Changing of the tank water can be conducted more often but must occur at least every 5 days.) Partial

changing of the tank water can be conducted within the 5-day period, if needed. The schedule of cleaning shall be constant across all test cohorts. If, for any reason, any one tank becomes soiled outside of the normal schedule and requires cleaning, this activity can be done without draining and cleaning the other tanks.

- Upon refilling of the tank, it shall be placed in its original position, and the camera shall be recalibrated to its original settings.

10.9.3. Experimental Environmental Cues

Visible Platform

- For nonspatial cued learning, minimizing visual spatial cues within the testing arena (enclosure) is critical to reduce the animal's access to visual cues that may be used to spatially navigate and to ensure the animal focuses on the platform as the cue.
- The tank shall be encircled by a curtained wall devoid of spatial cues.
- The experimenter shall remain at the start position during the trial to minimize movement or exit to the outside of the curtained wall.

Spatial Hidden Platform

- Visual cues within the testing room or within an enclosure are a major defining factor in the ability of an animal to learn the location of the platform via spatial processing.
- All tanks shall have an identical cue profile. If multiple testing areas are to be used, the environment shall be as identical as possible with regards to defined cues and architectural features.
- One specific cue set profile shall be used across all animals and shall be maintained over the entire course of the study.
- The rack of test animals within the test room is an odor and auditory cue. It can serve as a visual spatial cue, as well; thus, localization shall be standardized across test sessions and test rooms as dictated by room configuration. Location markers shall be employed as needed to ensure uniform rack placement across sessions.
- Room configuration. The room and walls shall be devoid of extraneous visual cues as much as possible, and any items remaining shall be considered part of the cue profile and thus, remain stationary across the study.
- Within a room, the background walls are light in color and, as an example, cues can consist of (1) two dark vertical lines from floor to ceiling (each line being 20 cm wide with a space of approximately 15 cm between them), (2) a large dark circle approximately 32 cm in diameter, and (3) a long dark horizontal line (20 cm in width) from the wall edge.
- Within-curtain enclosure. Curtains can be placed as walls around the tank to define a test enclosure. Curtains shall be placed a minimum of 2 ft. and a maximum of 3 ft. from the inner wall of the tank. The curtains shall be smooth and securely hung in a manner to minimize disturbance by testing room airflow. The curtains shall be at a sufficient distance from the tank to allow for experimenter mobility without movement of the curtain and disruption of the spatial cues. A permanent opening

shall be maintained for the experimenter to enter and exit and/or to remain (serving as a cue) during the trial to minimize any movement of the curtain and disruption of the spatial cues placed on curtain walls.

- The experimenter serves as a cue; thus, a mark on the floor shall be provided to indicate the exact location where the experimenter should remain after placing the animal into the water maze. The cues shall be placed within the animal's line of sight as it is swimming in the maze.
- A schematic of the room configuration or curtain dimensions and cues shall be included in the testing protocol (Figure 10-1).

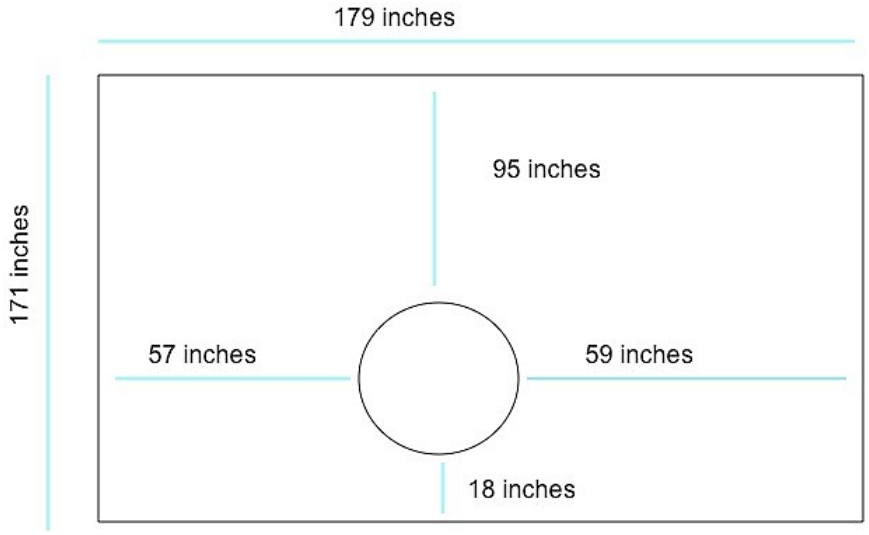

Figure 10-1. A Representative Example of Room Dimensions Relative to Water Tank

Lighting

The nature of lighting is critical to maximize accuracy of video tracking.

- Lighting shall be bright enough to allow for visualization of spatial cues.
- Lighting shall be bright enough to allow the video camera to track the animal.
- Lighting shall be arranged to prevent reflection on the water or in the video image (reflections can compromise video tracking as the software may confuse those with the animal).
- Lighting shall be arranged to prevent a shadow being cast into the tank interior from any surrounding structure, including the experimenter.
- Lighting shall be even, such as that obtained with a diffuse light source like a shaded fluorescent tube or globe-type incandescent bulb. Spotlights or uneven lighting shall not be used.
- Lighting shall be indirect and not in the direct line of sight of the camera. (One way this can be accomplished is by placing four to six globe bulbs around the pool, below the level of the water surface, outside of the line of sight of the camera lens.)

10.9.4. Configuration of Camera and Software

- The camera shall be positioned above the tank, perpendicular to the water surface. This position will be fixed and anchored above the tank. If this requires a stationary arm attached to the tank, that arm will be considered one of the visual spatial cues and thus, will be maintained in one position. Any wiring for the camera shall also be considered a spatial cue and be treated as such and included in the configuration schematic.
- The camera zoom setting shall be adjusted and the lens focused to display the entire experimental arena (in focus) on the computer screen.
- All camera automatic settings shall be disengaged.
- The camera aperture shall be adjusted for maximal contrast of the image.
- The entire setup shall be anchored during testing. The position of the arena, camera, and illumination shall be secured, relative to each other. If moved for tank cleaning, the system shall be recalibrated to its original position prior to testing.
- The camera zoom, focus control, and aperture setting shall be locked.
- Confirmation of the settings and quality of detection shall be conducted each day prior to the testing of the animals.
- The camera shall have a polarizing lens filter to minimize reflection.
- The visual field and lack of interference with capturing animals within that field (e.g., reflection) shall be confirmed each day prior to the start of the test session.
- Camera settings to capture the arena and animals shall be confirmed for alignment at the start of each test day.
- An automated video tracking system shall be used to capture MWM performance. The software algorithms used by the automated imaging program to define each of the endpoints shall be provided, per the commercial supplier's software manual. The software tracking system shall provide documentation of the testing protocol (arena parameters) and animal assignment and allow for post hoc evaluation of video images.

10.9.5. General Test Procedures

Placement and Removal of Animal from Tank

- The water shall be still before testing of any animal to minimize reflection that can interfere with video capture.
- The animal shall be removed from the transport cage and handled in a manner that provides support (e.g., placed in the crook of an arm, placed against body). With the animal supported by the palm of the hand or in a "carrier," it shall be placed into the tank by gently lowering the hand or carrier into the water. If an alternative to the experimenter's hand is used for animal placement into the tank or removal from the platform, this movement will be done in a manner that provides whole-body support to the animal and will be the method used for all animals in the study.
- The animal shall be placed in the tank with its nose facing the wall.

- If the animal cannot swim and sinks (not diving underwater), it shall be immediately removed from the tank and placed in the holding cage. For safety reasons, it is recommended that such retrieval be done with a container rather than by hand. This behavior shall be documented.
- Removal of animals from the platform shall be conducted with a whole-body method and not by use of the tail (unless necessary). The animal shall be first allowed to climb into the experimenter's hand or "carrier" prior to removal from the platform or from the water during the probe test. Upon removal, animals shall be placed on an absorbent towel. Effort shall be taken to place the animal in an area without air drafts to minimize discomfort. A "tested" animal shall not be placed into a cage with animals waiting for testing.

Start Location in the Tank

- A marking visible to the experimenter shall be placed on the outside of the tank to indicate starting location.
- Geographical nomenclature (N, NE, E, SE, S, SW, W, NW) is used to identify start locations (relative to cued/training goal/platform NE location) in the start location pattern provided. An alternative nomenclature can be used but must be consistent with the orientation and cross-referenced with geographical nomenclature.
- Cued learning (visible platform) start positions: SE to require a full transition across the tank to the platform in the NW quadrant.
- Different start quadrant locations will be randomized during spatial learning (hidden platform) and reversal learning. The sequences of start locations are designed to balance the right/left goal location (see Table 10-2 and Table 10-3 for example start positions).

Table 10-2. Start Locations for Acquisition with NE Quadrant as Hidden Platform Location

Day	Trial 1	Trial 2	Trial 3
1	S	W	NW
2	NW	S	SE
3	SE	NW	W
4	W	S	SE

Repeat sequence as needed.

Table 10-3. Start Locations for Reversal Learning with SW Quadrant as Hidden Platform Location

Day	Trial 1	Trial 2	Trial 3
1	N	E	SE
2	SE	N	NW
3	NW	SE	E
4	E	NW	N

Repeat sequence as needed.

Testing Schedule

- For each day of testing, a test schedule shall be provided, indicating the order that animals are to be tested and ensuring a counterbalance across dose groups. For any individual animal, this order shall be maintained across all test sessions.
- While animal performance will improve with training, the ITI between runs shall remain constant to maintain a uniform time for memory integration and learning.
- The testing schedule shall be appropriate for the species and age of the animal (e.g., adult rats perform well with consecutive training trials, whereas young rats and mice are more prone to hypothermia-induced performance effects; therefore, the latter's trials are separated by a significant ITI, approximately 30 minutes).

Visible Platform

- Animals will be trained on cued learning using a visible platform in the **NW** quadrant to ensure swimming ability, basic vision, and the ability to escape onto the platform.
- The task shall be performed under conditions that obscure the visibility of room spatial cues (e.g., decreased room lighting, curtain around tank with no cues attached).
- On day 1, the animal shall be placed on the platform for approximately 30 seconds to familiarize it with the platform before the initiation of the first trial.
- For each trial, the animal shall be placed into the periphery of the tank within a defined quadrant (Table 10-2), with its nose facing the wall of the tank.
- Animals will be allowed up to 90 seconds to find and escape onto the platform. If the animal does not find the platform within 90 seconds, the experimenter will gently guide the animal toward the platform by placing the hand behind the animal and allowing the animal to maintain a swimming action to reach the platform. The animal shall be allowed to escape the water by climbing onto the platform. This procedure will ensure that all animals learn to associate swimming and fully climbing onto the platform as the method of escape from the water. The animal will be allowed to remain on the platform for approximately 20 seconds. If the animal re-enters the water after climbing onto the platform, the same guidance procedure will be conducted, and the animal will be allowed to remain on the platform for approximately 10 seconds.
- Each animal will receive three sequential training trials per day for a total of 2 days with an ITI of at least 60 seconds for adult rats and a longer ITI for mice and young rats. The ITI length allows for integration of the learned event and minimizes the fatigue factor, thus decreasing variability in latency across trials. A relatively constant ITI will be maintained across the study for each animal in repeating the trial rotation for training.
- The start location sequence (SW, NE, SE) shall be followed in four sequential trials.
- If an animal "floats" or "circles" in early sessions, it may perform in later sessions and thus shall not be excluded. If an animal fails to swim (i.e., sinks but does not dive), this shall be noted. If this occurs on the second day, the animal shall be considered for removal from the study. Replacement of an animal at this point will

result in a lack of test history but shall be considered in discussions with the COR as an option to maintain sufficient n size.

- All animals shall achieve the performance criterion of reaching the visible platform before the 90-second cut-off before progressing to the hidden platform test. If an animal does not achieve the criterion within the number of sessions designated, a decision shall be made about additional training sessions.

10.9.6. Hidden Platform (Spatial Acquisition)

- Approximately 24 hours following completion of the visible platform test, the animal will be placed in the water maze and allowed 90 seconds to find the platform (NE quadrant). Animals shall remain on the platform for approximately 20 seconds, and they will be permitted to remain on the platform for approximately 20 seconds during all sessions of the first day and in the first trial of each subsequent day. In the second and third trial of subsequent testing days, the interval can be decreased to approximately 5 seconds.
- If the animal fails to find the platform within the maximum trial time, the experimenter will gently guide the animal toward the platform by placing the hand behind the animal and allowing the animal to maintain a swimming action to reach the platform. The animal shall be allowed to escape the water by climbing onto the platform. If the animal re-enters the water from the platform, the same guidance procedure will be conducted, and the animal allowed to remain on the platform for approximately 10 seconds. This procedure will ensure all animals learn to associate swimming and fully climbing onto the platform as the method of escape from the water as well as visual cues for spatial orientation to the platform. The animal will not be picked up from the water and placed on the platform. The animal will be allowed to remain on the platform for approximately 20 seconds.
- Over a 7-day interval, each animal will receive three sequential training trials per day with an ITI of at least 60 seconds (time constant across animals) for rats, with a longer ITI for mice or young rats (due to fatigue).
- It is expected that approximately 80% of control animals will reach the criterion of ≥50% decrease in either latency to platform or swimming distance to platform by the seventh day. If not, considerations shall be made to extend the hidden platform spatial learning phase of the assay.

10.9.7. Probe Trial (Reference Memory)

The MWM probe trial allows for the confirmation and assessment of spatial reference memory in performance of the task.

- At 24 ± 2 hours following each animal's last hidden platform test (acquisition), the animal will be assessed for reference memory using a single probe trial.
- Visual cues and lighting conditions shall remain as they were for the hidden platform task, but the platform will be removed from the tank.
- As described in the hidden platform testing, the animal shall be placed in the tank at the SW start location.

- The animal shall be allowed to freely swim for 90 seconds with data collected for each 30-second epoch. The animal shall be removed from the tank at the end of 90 seconds.

10.9.8. Reverse Platform (Reversal Learning)

Approximately 48 ± 4 hours after the probe trial (scheduling to maintain constant time interval across all groups), the animals will be evaluated for their ability to learn a new platform location.

- The tank setup and visual cues shall be the same as those for the hidden platform test, with the platform moved to the opposite quadrant location (**SW**).
- Animals will be individually placed in the water maze and allowed up to 90 seconds to find the platform. If an animal fails to find the platform, it will be guided to and placed upon the platform. Animals will be allowed to remain on the platform for approximately 20 seconds.
- Given that the animals have already learned the parameters of the task and are only shifting location, they shall be tested for three trials per day for 3 days with a minimal ITI of 60 seconds.
- If the performance fails to reach the level observed on the last day of the hidden platform acquisition (the last day prior to the probe trial), training shall continue for up to 2 additional days or until performance reaches the previous levels observed in the hidden platform test or at the decision of the COR.

10.9.9. Endpoints Collected

For the MWM, data collected for visible platform, hidden platform, and reversal learning shall be averaged over the daily trials for each animal for repeated measures analysis to demonstrate acquisition and, for the reversal learning, the ability to shift to a new location. For the probe test, data are collected to demonstrate preference for the quadrant previously containing the escape platform, relative to the other quadrants. It is also used to show that the animal can learn that the platform is not present and then shift their search strategy.

Visible Platform (Nonspatial Learning)

- Time to find platform (latency)
- Total distance traveled to platform (path length)
- Average swim speed
- Time spent floating (% trial duration)
- Percent thigmotaxis time (% trial duration when the subject was in the outer 10% of the pool diameter)
- Thigmotaxic tendency (proportional distance traveled within the outer 10% of the pool relative to total distance traveled)
- Daily averages calculated for trials within a session for individual animals
- Analysis: RMANOVA (dose and day as factors)

Hidden Platform (Spatial Acquisition)

- Time to reach platform (latency)
- Total distance to reach platform (path length)
- Average swim speed
- Time spent floating (% trial duration)
- Percent thigmotaxis time as % trial duration when the subject was in the outer 10% of the pool diameter
- Pathway tracking, as available, by commercial video tracking and analysis software, as instructed by manufacturer
- Daily averages calculated for trials within a session for individual animals
- Analysis: RMANOVA (dose and day as factors)

Probe Trial (Reference Memory)

Initial Response

- Initial latency to enter the quadrant containing previous platform location
- Initial latency to swim to a predefined annulus surrounding the previous platform target that is 1.5 times larger than the target itself
- Initial latency to swim to previous platform target site
- Distance traveled to the entry into the previous target platform quadrant
- Distance traveled to the target annulus
- Distance traveled to the previous platform target site
- Data analyzed by one-way ANOVA (dose as factor; or Kruskal Wallis/Dunn)

Total Session and 30-second Epochs

- Platform-site crossings: number of crossings over the previous escape platform location
- Platform-annulus site crossings: number of crossings over the annulus of the previous escape platform location
- Time in the target annulus site
- Total number of entries into each quadrant
- Quadrant time: the total time spent in each quadrant
- Quadrant distance traveled: the total swimming distance (path length) within each quadrant
- Calculated quadrant time percentage: the percentage of time spent in each quadrant
- Calculated quadrant distance traveled percentage: the percentage of distance traveled in each quadrant
- Total session swimming distance: the total swimming distance (path length) covered over the entire tank over the full session

- Search strategy: pathway tracing as provided by instrument manufacturer software analysis
- Data analyzed by one-way ANOVA (dose as factor, or Kruskal Wallis/Dunn): platform-site crossings, platform-annulus site crossings, time in target annulus site, in-goal quadrant-only total time spent; total distance traveled, number of entries; calculated % total time spent, % total distance traveled in-goal quadrant relative to other quadrants to demonstrate preference for the goal quadrant
- Analyses of time spent, distance traveled, or entries into quadrants other than goal quadrant are not of relevance to the assessment and shall not be conducted

Reversal Platform (Reversal Learning)

- Time to reach platform (latency)
- Total distance traveled to reach platform (path length)
- Average swim speed
- Time spent floating (% trial duration)
- Calculated percent thigmotaxis time as % trial duration when the subject was in the outer 10% of the pool diameter
- Pathway tracking by commercial video tracking and analysis software, as instructed by the manufacturer, shall be maintained for all MWM trials

10.9.10. Statistical Analysis

- Data shall be evaluated for homogeneity of variance and normally distributed observations. Transformation of data or nonparametric methods should be considered if these assumptions are violated.
- Independent group mean comparisons shall be conducted upon significant one-way ANOVA or Kruskal Wallis/Dunn results.
- Post hoc analysis shall be conducted in the presence of significant main effects and the absence of significant interactions.

10.10. References

Baarendse PJJ, Counotte DS, O'Donnell P, Vanderschuren LJMJ. 2013. Early social experience is critical for the development of cognitive control and dopamine modulation of prefrontal cortex function. Neuropsychopharmacology. 38(8):1485-1494. https://doi.org/10.1038/npp.2013.47

Bailey KR, Rustay NR, Crawley JN. 2006. Behavioral phenotyping of transgenic and knockout mice: Practical concerns and potential pitfalls. ILAR J. 47(2):124-131. https://doi.org/10.1093/ilar.47.2.124

Bohlen M, Cameron A, Metten P, Crabbe JC, Wahlsten D. 2009. Calibration of rotational acceleration for the rotarod test of rodent motor coordination. J Neurosci Methods. 178(1):10-14. https://doi.org/10.1016/j.jneumeth.2008.11.001

Burke AR, McCormick CM, Pellis SM, Lukkes JL. 2017. Impact of adolescent social experiences on behavior and neural circuits implicated in mental illnesses. Neurosci Biobehav Rev. 76(Pt B):280-300. https://doi.org/10.1016/j.neubiorev.2017.01.018

Chapillon P, Lalonde R, Jones N, Caston J. 1998. Early development of synchronized walking on the rotorod in rats: Effects of training and handling. Behav Brain Res. 93(1-2):77-81. https://doi.org/10.1016/s0166-4328(97)00137-x

Crabbe JC, Wahlsten D, Dudek BC. 1999. Genetics of mouse behavior: Interactions with laboratory environment. Science. 284(5420):1670-1672. https://doi.org/10.1126/science.284.5420.1670

Crawley JN. 1999. Behavioral phenotyping of transgenic and knockout mice: Experimental design and evaluation of general health, sensory functions, motor abilities, and specific behavioral tests. Brain Res. 835(1):18-26. https://doi.org/10.1016/s0006-8993(98)01258-x

Csomor PA, Yee BK, Vollenweider FX, Feldon J, Nicolet T, Quednow BB. 2008. On the influence of baseline startle reactivity on the indexation of prepulse inhibition. Behav Neurosci. 122(4):885-900. https://doi.org/10.1037/0735-7044.122.4.885

Gallagher M, Burwell R, Burchinal M. 1993. Severity of spatial learning impairment in aging: Development of a learning index for performance in the Morris water maze. Behav Neurosci. 107(4):618-626. https://doi.org/10.1037//0735-7044.107.4.618

Genzel L. 2021. How to control behavioral studies for rodents—Don't project human thoughts onto them. eNeuro. 8(1):ENEURO.0456-0420.2021. https://doi.org/10.1523/eneuro.0456-20.2021

Gerlai R. 2001. Behavioral tests of hippocampal function: Simple paradigms complex problems. Behav Brain Res. 125(1-2):269-277. https://doi.org/10.1016/s0166-4328(01)00296-0

Geyer MA, Swerdlow NR. 1998. Measurement of startle response, prepulse inhibition, and habituation. Curr Protoc Neurosci. 3(1):8.7.1-8.7.15. https://doi.org/10.1002/0471142301.ns0807s03

Gouveia K, Hurst JL. 2017. Optimising reliability of mouse performance in behavioural testing: The major role of non-aversive handling. Sci Rep. 7:44999. https://doi.org/10.1038/srep44999

Graham DL, Meyer JS, Stanwood GD. 2018. Chapter 25 - Behavioral phenotyping in developmental neurotoxicology—Simple approaches using unconditioned behaviors in rodents. In: Slikker W, Paule MG, Wang C, editors. Handbook of Developmental Neurotoxicology. 2nd ed. London, UK: Academic Press. p. 287-308.

Hånell A, Marklund N. 2014. Structured evaluation of rodent behavioral tests used in drug discovery research. Front Behav Neurosci. 8:252. https://doi.org/10.3389/fnbeh.2014.00252

Harry GJ, McBride S, Witchey SK, Mhaouty-Kodja S, Trembleau A, Bridge M, Bencsik A. 2022. Roadbumps at the crossroads of integrating behavioral and in vitro approaches for neurotoxicity assessment. Front Toxicol. 4:812863. https://doi.org/10.3389/ftox.2022.812863

Holson RR, Freshwater L, Maurissen JPJ, Moser VC, Phang W. 2008. Statistical issues and techniques appropriate for developmental neurotoxicity testing: A report from the ILSI Research Foundation/Risk Science Institute expert working group on neurodevelopmental endpoints. Neurotoxicol Teratol. 30(4):326-348. https://doi.org/10.1016/j.ntt.2007.06.001

Hormigo S, Cardoso A, Sancho C, López DE, Moreno C. 2019. Associations between sensorimotor gating mechanisms and athletic performance in a variety of physical conditioning tests. Eur J Appl Physiol. 119(4):921-932. https://doi.org/10.1007/s00421-019-04081-1

International Programme on Chemical Safety (IPCS). 1986. Environmental Health Criteria 60. Principles and methods for the assessment of neurotoxicity associated with exposure to chemicals. Geneva, Switzerland: World Health Organization. https://apps.who.int/iris/handle/10665/40136. [Accessed: October 29, 2022]

Kafkafi N, Benjamini Y, Sakov A, Elmer GI, Golani I. 2005. Genotype-environment interactions in mouse behavior: A way out of the problem. Proc Natl Acad Sci U S A. 102(12):4619-4624. https://doi.org/10.1073/pnas.0409554102

Kafkafi N, Agassi J, Chesler EJ, Crabbe JC, Crusio WE, Eilam D, Gerlai R, Golani I, Gomez-Marin A, Heller R et al. 2018. Reproducibility and replicability of rodent phenotyping in preclinical studies. Neurosci Biobehav Rev. 87:218-232. https://doi.org/10.1016/j.neubiorev.2018.01.003

Lampe JF, Ruchti S, Burman O, Würbel H, Melotti L. 2019. Play like me: Similarity in playfulness promotes social play. PLoS One. 14(10):e0224282. https://doi.org/10.1371/journal.pone.0224282

Lewejohann L, Reinhard C, Schrewe A, Brandewiede J, Haemisch A, Görtz N, Schachner M, Sachser N. 2006. Environmental bias? Effects of housing conditions, laboratory environment and experimenter on behavioral tests. Genes Brain Behav. 5(1):64-72. https://doi.org/10.1111/j.1601-183X.2005.00140.x

Maei HR, Zaslavsky K, Teixeira CM, Frankland PW. 2009. What is the most sensitive measure of water maze probe test performance? Front Integr Neurosci. 3:4. https://doi.org/10.3389/neuro.07.004.2009

Mandillo S, Tucci V, Hölter SM, Meziane H, Banchaabouchi MA, Kallnik M, Lad HV, Nolan PM, Ouagazzal AM, Coghill EL et al. 2008. Reliability, robustness, and reproducibility in mouse

behavioral phenotyping: A cross-laboratory study. Physiol Genomics. 34(3):243-255. https://doi.org/10.1152/physiolgenomics.90207.2008

Marcotte M, Bernardo A, Linga N, Pérez-Romero CA, Guillou JL, Sibille E, Prevot TD. 2021. Handling techniques to reduce stress in mice. J Vis Exp. (175):e62593. https://doi.org/10.3791/62593

Maurissen JPJ, Marable BR, Andrus AK, Stebbins KE. 2003. Factors affecting grip strength testing. Neurotoxicol Teratol. 25(5):543-553. https://doi.org/10.1016/s0892-0362(03)00073-4

Meyer OA, Tilson HA, Byrd WC, Riley MT. 1979. A method for the routine assessment of fore- and hindlimb grip strength of rats and mice. Neurobehav Toxicol. 1(3):233-236.

Miller EA, Kastner DB, Grzybowski MN, Dwinell MR, Geurts AM, Frank LM. 2021. Robust and replicable measurement for prepulse inhibition of the acoustic startle response. Mol Psychiatry. 26(6):1909-1927. https://doi.org/10.1038/s41380-020-0703-y

Moore DS, McCabe GP. 2001. Introduction to the practice of statistics. 3rd ed. New York, NY: W.H. Freeman and Company.

Moser VC. 2011. Functional assays for neurotoxicity testing. Toxicol Pathol. 39(1):36-45. https://doi.org/10.1177/0192623310385255

NAFTA Technical Working Group on Pesticides (NAFTA-TWG). 2016. Developmental neurotoxicity study guidance document. North American Free Trade Agreement. https://www.epa.gov/sites/default/files/2017-02/documents/developmental_neurotoxicity_study_internal_guidance_document_final_0.pdf. [Accessed: October 29, 2022]

Oliveira AFS, Rossi AO, Silva LFR, Lau MC, Barreto RE. 2010. Play behaviour in nonhuman animals and the animal welfare issue. J Ethol. 28:1-5. https://doi.org/10.1007/s10164-009-0167-7

Panksepp J, Beatty WW. 1980. Social deprivation and play in rats. Behav Neural Biol. 30(2):197-206. https://doi.org/10.1016/s0163-1047(80)91077-8

Pellis SM, Pellis VC. 1990. Differential rates of attack, defense, and counterattack during the developmental decrease in play fighting by male and female rats. Dev Psychobiol. 23(3):215-231. https://doi.org/10.1002/dev.420230303

Pellis SM, Pellis VC, Burke CJ, Stark RA, Ham JR, Euston DR, Achterberg EJM. 2022. Measuring play fighting in rats: A multilayered approach. Curr Protoc. 2(1):e337. https://doi.org/10.1002/cpz1.337

Pilz PKD, Schnitzler HU. 1996. Habituation and sensitization of the acoustic startle response in rats: Amplitude, threshold, and latency measures. Neurobiol Learn Mem. 66(1):67-79. https://doi.org/10.1006/nlme.1996.0044

Pilz PKD, Arnold SW, Rischawy AT, Plappert CF. 2014. Longterm-habituation of the startle response in mice is stimulus modality, but not context specific. Front Integr Neurosci. 7:103. https://doi.org/10.3389/fnint.2013.00103

Pisula W, Modlinska K. 2020. Protocol for measuring free (low-stress) exploration in rats. Bio Protoc. 10(2):e3485. https://doi.org/10.21769/BioProtoc.3485

Rustay NR, Wahlsten D, Crabbe JC. 2003a. Influence of task parameters on rotarod performance and sensitivity to ethanol in mice. Behav Brain Res. 141(2):237-249. https://doi.org/10.1016/s0166-4328(02)00376-5

Rustay NR, Wahlsten D, Crabbe JC. 2003b. Assessment of genetic susceptibility to ethanol intoxication in mice. Proc Natl Acad Sci U S A. 100(5):2917-2922. https://doi.org/10.1073/pnas.0437273100

Saré RM, Lemons A, Smith CB. 2021. Behavior testing in rodents: Highlighting potential confounds affecting variability and reproducibility. Brain Sci. 11(4):522. https://doi.org/10.3390/brainsci11040522

Sensini F, Inta D, Palme R, Brandwein C, Pfeiffer N, Riva MA, Gass P, Mallien AS. 2020. The impact of handling technique and handling frequency on laboratory mouse welfare is sex-specific. Sci Rep. 10(1):17281. https://doi.org/10.1038/s41598-020-74279-3

Shoji H, Miyakawa T. 2018. Relationships between the acoustic startle response and prepulse inhibition in C57BL/6J mice: A large-scale meta-analytic study. Mol Brain. 11(1):42. https://doi.org/10.1186/s13041-018-0382-7

Slikker W, Jr., Acuff K, Boyes WK, Chelonis J, Crofton KM, Dearlove GE, Li A, Moser VC, Newland C, Rossi J et al. 2005. Behavioral test methods workshop. Neurotoxicol Teratol. 27(3):417-427. https://doi.org/10.1016/j.ntt.2005.02.003

Stark R, Pellis SM. 2020. Male Long Evans rats reared with a Fischer-344 peer during the juvenile period show deficits in social competency: A role for play. Int J Play. 9(1):76-91. https://doi.org/10.1080/21594937.2020.1720142

Takeshita H, Yamamoto K, Nozato S, Inagaki T, Tsuchimochi H, Shirai M, Yamamoto R, Imaizumi Y, Hongyo K, Yokoyama S et al. 2017. Modified forelimb grip strength test detects aging-associated physiological decline in skeletal muscle function in male mice. Sci Rep. 7:42323. https://doi.org/10.1038/srep42323

Valsamis B, Schmid S. 2011. Habituation and prepulse inhibition of acoustic startle in rodents. J Vis Exp. (55):e3446. https://doi.org/10.3791/3446

Vanderschuren LJMJ, Achterberg EJM, Trezza V. 2016. The neurobiology of social play and its rewarding value in rats. Neurosci Biobehav Rev. 70:86-105. https://doi.org/10.1016/j.neubiorev.2016.07.025

VanRyzin JW, Marquardt AE, McCarthy MM. 2020. Assessing rough-and-tumble play behavior in juvenile rats. Bio Protoc. 10(1):e3481. https://doi.org/10.21769/BioProtoc.3481

Vorhees CV, Williams MT. 2006. Morris water maze: Procedures for assessing spatial and related forms of learning and memory. Nat Protoc. 1(2):848-858. https://doi.org/10.1038/nprot.2006.116

Vorhees CV, Williams MT. 2021. Issues in the design, analysis, and application of rodent developmental neurotoxicology studies. Neurotoxicol Teratol. 87:107018. https://doi.org/10.1016/j.ntt.2021.107018

Wahlsten D. 2011. Mouse behavioral testing: How to use mice in behavioral neuroscience. Amsterdam: Academic Press.

10.11. Peer Review

The Division of Translational Toxicology (DTT) conducted a peer review of chapter 10 within the draft *Specifications for the Conduct of Toxicity Studies by the Division of Translational Toxicology at the National Institute of Environmental Health Sciences* by letter in February 2022 by the expert listed below. Reviewer selection and document review followed established DTT practices. The reviewer was charged to:

1. Peer review the following chapter within the draft Specifications for the Conduct of Toxicity Studies by the Division of Translational Toxicology at the National Institute of Environmental Health Sciences.
    - Chapter 10: Neurobehavioral Testing
2. Comment on the completeness of each chapter.

DTT carefully considered reviewer comments in finalizing this document.

Peer Reviewer

Kevin Crofton, Ph.D.
President
R3Fellows LCC
Durham, North Carolina, USA