
Methods for Assessing Neurodevelopmental Disorders in Rodents: A Critical Review of Behavioral Tests and Methodological Considerations Searching to Improve Reliability

[Boniface Ezenwata Echefu](#) , [Maria Becker](#) , Dan Stein , [Asher Ornoy](#) *

Posted Date: 3 February 2025

doi: 10.20944/preprints202502.0040.v1

Keywords: neurobehavioral disorders; behavioral tests; rodents; pitfalls; scoring system



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Review

Methods for Assessing Neurodevelopmental Disorders in Rodents: A Critical Review of Behavioral Tests and Methodological Considerations Searching to Improve Reliability

Boniface Echefu ¹, Maria Becker ¹, Dan Stein ¹ and Asher Ornoy ^{1,2,3,*}

¹ Department of Morphological Sciences and Teratology, Adelson School of Medicine, Ariel University, Ariel 40700, Israel

² Hadassah Academic College Jerusalem

³ Hebrew University Hadassah Medical School, Jerusalem 9112102, Israel

* Correspondence: ashero@ariel.ac.il or ornoy@cc.huji.ac.il

Abstract: Many neurobehavioral tests are used for the assessment of human-like behaviors in animals. Most of them were developed in rodents and are used for the assessment of animal models that mimic human neurodevelopmental and neuropsychiatric disorders (NDD). We have described tests for assessing social behavior, social interaction and social communication; tests for restricted and repetitive behaviors; tests for cognitive impairment, for sensory stimuli, for anxiety like behavior and for motor coordination deviations. These tests are used to demonstrate autistic-like behavior as well as other NDD. We described possible general pitfalls in the performance of such studies, as well as probable individual errors for each group of tests assessing specific behavior. The mentioned pitfalls may induce crucial errors in the interpretation of the results, minimizing the reliability of specific models of defined human NDD. It is imperative to minimize these pitfalls and also use sufficient and reliable tests that are able to demonstrate as many of the traits of the human disorder, grade the severity of the specific deviations and the severity of the tested NDD by using a scoring system. Due to possible gender differences in the clinical presentations of NDD, it is important to carry out studies on males and females.

Keywords: neurobehavioral disorders; behavioral tests; rodents; pitfalls; scoring system

1. Introduction

Neurodevelopmental and neurobehavioral disorders (NDD) such as attention deficit hyperactivity disorders (ADHD), obsessive compulsive disorders (OCD), autism spectrum disorders (ASD) are highly prevalent in humans [1]. Other conditions, including intellectual disability, bipolar disorder, Angelman syndrome, Tourette syndrome (TS) or schizophrenia (ZF), are associated with disruptions in the intricate sequence of brain developmental processes. These processes include proliferation, differentiation, cellular migration, synapse formation, and abnormalities in neural network connectivity [2–4]. The etiology and pathophysiology of NDDs are generally attributed to a combination of biological, genetic, epigenetic, and environmental factors[5,6].

Laboratory animals, particularly rodent models, remain crucial tools for face validation and predictive treatment developments for human NDD as well as other diseases. Studies employing rodents to investigate behavioral changes analogous to symptoms inherent in human NDD often rely on several standardized neurobehavioral and neurological test paradigms. For several decades, these tests have been extensively utilized to understand behavioral phenotypes in mammals contributing significantly to advancements in treatment development for human disorders, including NDDs. Nevertheless, these paradigms are not without challenges, limitations and drawbacks. A major issue

is the frequent failure to replicate results, which can often be attributed to variations across laboratories, procedural modifications, inconsistencies in data collection and presentation, litter effects, and the technical expertise of personnel, which collectively pose significant challenges to the reliability and generalizability of findings. Efforts to improve standardization, enhance reproducibility, and account for these variables remain critical to advancing the translational relevance of rodent models in NDD research. Many neurodevelopmental and neurobehavioral disorders in human are diagnosed using validated developmental and behavioral tools [1]. The American Psychiatric Association emphasizes the use of standardized scoring system and the demonstration of characteristic behavioral changes that reflect specific disorders. Typically, diagnosis involves the use of well-validated questionnaires and behavioral tests complemented by a detailed history from the individual or caregivers. The behavior of the individual is assessed across multiple domains and systematically scored to ensure accuracy and reliability in diagnosis. Similarly, in animal models of neurodevelopmental disorders, achieving a proper and specific diagnosis requires validating a set of observable behavioral deviations and classifying the degree of their severity. This process involves the development of robust and reproducible scoring systems that align with established human diagnostic criteria, where applicable. Such classification is essential not only for understanding the behavioral phenotypes in animal models but also for ensuring their translational relevance in preclinical research [5].

One of the major challenges in diagnosing NDD is the lack of specific biological markers for these disorders. This limitation extends to animal models, where researchers must primarily rely on behavioral tests that depict particular behavioral changes. Moreover, even when observing behavioral traits that closely mimic human normal or abnormal behaviors, grading the severity of these changes remains challenging, as it often falls short of the precision required for human diagnosis.

In this review we will discuss the most reliable and widely used tests in rodent models for identifying behavioral deviations that resemble those observed in various NDDs. Furthermore, we will highlight the potential pitfalls and limitations in the evaluation and interpretation of these test results. Addressing these challenges is critical for enhancing the validity and reliability of rodent model data, ultimately improving the translational value of such studies for human clinical applications.

2. Possible Pitfalls When Using Animal Models of NDD: Potential General Problems

2.1. Litter Effect

Litter effects significantly impact embryonic and early postnatal studies due to the near impossibility of timing conception accurately [7]. Although awareness of litter effects remains relatively low, its effects are responsible for 30–60% of the variability commonly encountered in behavioral studies that tend to persist into adult life [8]. Studies focusing on early developmental stages should prioritize controlling for these effects to avoid misleading conclusions from small litter numbers or disproportionate litter size [7]. Despite the existence of literature highlighting the importance of litter control, many researchers preferred resources conservation practice using smaller underpowered samples. These practices end up leading to increased false positive rates and introduce artefacts that can obscure genuine treatment or genetic effects in test animals [9,10]. To address these issues, researchers should not only increase the number of litters to achieve higher statistical power but also equalize litter sizes so that all litters have uniform numbers starting on post-natal day (PND) one. For example, in the valproic acid (VPA) model of autism, administering VPA to pregnant females induces autism-like phenotypes in their offspring [136]. In this type of experimental design, the sample size should be determined by the number of litters in addition to the total number of offspring. Failing to account for litter effects in this way may result in a significantly biased and severely underpowered outcome. Implementing appropriate measures to

mitigate this significant source of variability is crucial for improving the rigor, accuracy, and reproducibility of experimental results, particularly in studies assessing NDDs. Proper control of litter effects will not only enhance the reliability of rodent model research but also strengthen the translational value of such studies for human clinical applications.

2.2. Uniformity of Pregnancy Timing

The importance of timing accuracy for all procedures starting from mating, vaginal plug observation, treatment, parturition up to neurobehavioral testing sessions in studies of NDDs cannot be overemphasized. Considering the dramatic changes during the short estrous cycle and rapid rate of development in rodents' central nervous system (CNS), mating hours and subsequent confirmation of copulation should be precise and strictly adhered to. Proper application of the Whitten and Lee-boot effect in study design is strongly encouraged for the purpose of synchronizing and maintaining reproductive events, that notwithstanding, male to female mating ratio should not exceed 1:3 [11,12]. In studies that would involve administration/exposure of dams to chemical/biological manipulations, it is fundamental to ensure that all dams undergo a similar procedure (including handling) devoid of time-gap and stress [13,14]. Time disparity could mean significant changes in studies that model environmental components of diseases such as VPA in ASD because of the dramatic changes per time in rodent embryonic life [15]. Although rodents are nocturnal animals, many researchers perform behavioral assays during the day. Consequently, some studies have been conducted in night/darkness in attempts to understand the significance of circadian cycle on behavior and how it affects results of studies conducted in the active and resting phases in rodents [16] and that may be crucial for performing behavioral tests. Regardless of the phases of rodent activity, we opine that conducting tests at the same time of the day, having minimal number of litters (at most 2) or few animals (at most 20) per day, equal resting and inter-test intervals for all litters experimental groups are crucial to circumventing challenges related to time optimization in rodent behavioral test. This is evidently plausible as behaviors in rodents are traditionally measured in time units.

2.3. Housing and Testing Conditions

It is critical to briefly lay out some of the factors that could interfere with study outcomes of NDDs leading to conflicting results that are usually encountered between laboratories and experimental cohorts. Housing and testing conditions, as basic as they appear in animal studies, play a significant role in the phenotypes of rodent models of NDDs [17]. Studies in uncompartimentalized pre-weaning rodents showed that the males tend to be more social, compared to males separated early from the females, and that estrous phase of females tends to influence their social behavioral outcomes and other sensory-motor reflexes [18,19]. It is recommended therefore to employ Lee-Boot effect where applicable for maintenance of steady phase of estrus during behavioral studies especially for ASD models. Studies have reported that mice exhibit defensive behaviors when housed in proximity to rats. Therefore, it is advisable to avoid housing these species as neighbors. Additionally, when common equipment is used for both species, thorough cleaning is essential to minimize stress and cross-species influences [20,21]. Lighting conditions, temperature and humidity must be tightly controlled and well documented for dams and their offspring at all times [22,23]. Strong strange smell and sound can be sources of stress for both dam and offspring and might impact the results of several types of neurobehavioral assays in high magnitude [13,14]. Cage and bedding replacement should be avoided at least two days prior to behavioral testing. Reports have emerged that rodents display several hours of increased anxiety, motor activity and decreased sleep time on the days that the bedding materials and cages were changed [19,24–26]. Familiarization and consistency of experimenter in terms of scent, handling and treatment techniques play a vital role in maintaining evenness of behavior in rodent models. When selecting a control group, it is important to recognize that the behavior of vehicle-control animals may differ from that of normal controls. This is because the vehicle itself can sometimes alter specific behavioral phenotypes, potentially affecting the study's

results [20] While considering that some of the factors highlighted so far can be significantly influenced by genetic components of the test animals, it is also important to note that the test-order, test-number and the intervals between tests might also be sources of conflicting results in neurobehavioral studies.

2.4. *Phenotyping Specificity of Test Paradigms*

Specificity of test paradigm relative to different behavioral phenotype poses a complex burden on choosing suitable test battery for studying a given NDD such as ASD. Similarly, different tests are currently available for several behavioral symptoms that characterize a specifically known disorder [17]. Regardless of the fundamental importance of these issues, less effort has been made to resolve them, leading to scarcity and paucity in literature. Ornoy et al. have pioneered the movement calling for a scoring system for ASD studies in an attempt to solve the problem of disease-model misidentification [5]. The study's scope heralds the need for implementation of a comprehensive scoring system for accurate face and construct validity of animal models, and recapitulation of human ASD features as a prerequisite for the translation of preclinical findings into suitable clinical tools. Another compounding challenge is the contrasting outcome obtained when comparing behavioral test results from infancy, adolescence and adult stage of rodents of NDDs [27]. Life stage specific disparity could be minimized by starting behavioral tests early during infancy and progressing through adolescence to adult life, so that vital information that are developmental-stage specific are well captured [28].

Attempts to broaden the scope shows that a handful of studies are increasingly reporting on the commonality of phenotypic dysfunction and mechanisms shared among NDDs especially SZ, ASD and ADHD [17,29–32]. Considering these reports, the use of only few tests to characterize a model of NDD will not yield a plausible conclusion on the specificity of that model. We believe that a model can be termed broadly as neurodevelopmental-model or more specifically as model of asocial/repetitive/communication behavior without necessarily ascribing it to a particular named NDD (e.g., SZ, OCD). This popular practice of hasty mischaracterization of rodent models of NDD is a major source of conflicting results in the field of NDD research. Such practices negatively impact the integrity of findings and hamper the translation of animal findings to human clinical use. Applying the proposed scoring system would help to identify not only ASD models but also other NDDs models for the evaluation and improvement of preventive and therapeutic interventions [5]. Regardless of the heterogeneous features of NDDs, high co-occurrence and arbitrary nature of the boundaries that define individual disorders, having a comprehensive study design and robustized test battery remain the crucial ingredients of obtaining reliable and reproducible neurobehavioral assay results.

3. Behavioral Tests Assessing Specific Behavioral Traits

3.1. *Tests for Social Behaviors*

3.1.1. The Three-Chamber Test for Sociability and Novel Social Preference

In order to assess socialization and social preference in preclinical animal investigations, the three chambers (3-chamber) test is unarguably the most globally preferred behavioral paradigm as adopted from Crawley et-al[5]. The test is classically employed to test sociability index in rodent studies. In addition, the test has been upgraded for the concomitant observation of preference in social novelty. Basically, sociability is defined in terms of the time a given subject spends with an age and sex-matched conspecific. The apparatus (arena) consisted of a rectangular box partitioned into three separate chambers. The two outer chambers house two circular wire cages and are opened to the middle chamber via 2 square doors. The color and size of the apparatus and cages vary. Despite the consistent principles and basic design, significant methodological variations exist across research laboratories, including differences in size, cage materials, test duration, room acclimatization periods, and behavioral outcome interpretation criteria.

In the test's original version, subjects are introduced into the arena usually into the center of the middle chamber for habituation in the presence of the circular empty cages at opposite corners of the outer chambers. 5 to 10 minutes may be allocated for this stage, though it is often skipped entirely. In the second phase, a conspecific is placed in one of the circular cages and the subject is allowed to explore the arena once again for 10 minutes. The objective of these two phases is to assess sociability index in the subject [33–35]. The modified version is aimed at observing social preference/memory, is achieved by incorporating a third phase to the test, during which a second conspecific is placed in the remaining empty circular cage. Once again, the test subject allows 10 minutes to explore the arena with the first and second conspecific in the two circular cages. The time spent in each chamber, the time spent around each cage and the total distances moved by test subject are measured [36]. According to the principle of the test, animals that spend equal or more significant time with the empty cage than with the first conspecific in the second phase are considered to have deficits in social interaction. Similarly, subjects that spend significant time with the first familiar conspecific than with the second conspecific in the modified version is adjudged to have impaired social memory and lack interest in social novelty [37,38].

3.1.2. The Partition Test for Abnormalities in Social Behavior

The partition test, much like 3-chamber test, is employed to assess social behavioral abnormality in rodents and other animal models of autistic-like behavior [5]. This test is conducted preferably in a standard home cage. Nevertheless, standard apparatus is routinely used especially when the system is automated [39]. As the name applies, the cage is partitioned into two equal halves using a perforated wall. The wall can be transparent or opaque but allows sensory contact between a test animal and a conspecific (stranger). The test involves 2 phases of habituation and task performance. In the first phase, the test subject is introduced to one part of the partition and allowed variable time (5 or 10 minutes) to explore and get familiar with the empty arena. This is followed by the second phase which involves introduction of a conspecific to the other part of the partition and once again the test subject is allowed 10 minutes to explore the conspecific. The time spent by the test animal near the partition in the second phase compared to opposing end of the arena measures sociability index [36]. A test subject that fails to spend significantly more time near the partition with the stranger is considered to be socially impaired [39,40]. Additionally, attempts by the test animal to put nose/paws into the perforations or try to gnaw at the partition could be considered an act of nonverbal communication [41,42]. Changing the stranger with other conspecifics could enable the test to assay for novel and social preference in the animal [36,40,43].

3.1.3. Reciprocal Social Interactions

This assay paradigm investigates the social behaviors exhibited by animals in response to interactions with a conspecific, providing insights into their social dynamics and potential behavioral anomalies [44–46]. The apparatus could be a common home cage as the test is best suited for manual observation [47]. In the first step, a test animal is introduced and allowed to habituate the test arena for 5 minutes followed by introduction of a conspecific with which it will be allowed to interact for another 10 minutes [47]. An observer measures the duration and frequency by which the test animal initiates actions (involving sniffing, following, chasing, pushing past, crawling over/under, pushing under, grooming, fast-paced wrestling, pouncing, spinning, chasing, biting and boxings [48,49]. The primary aim is to assess the degree of social engagement, whereby failure to interact more frequently or reciprocate behaviors is indicative of asocial traits akin to NDDs [50,51].

3.1.4. Social Conditioned-Place Preference and Motivation

This test is primarily used to evaluate how rewarding a test animal finds a socially conditioned environment [51]. The test set up consists of 2-chambers with each chamber's wall differently marked and one of the chambers housing a social stimulus. Briefly test subject is placed in the first chamber

that has special markings on the walls that will serve as cue as well as a social stimulus (like conspecific or pheromone). In the second phase the subject is again placed in a new chamber with different wall markings but lacking the social stimuli. In the final stage, subject is introduced into the arena with the two chambers opened to each other but without the social stimulus in either of the chambers [44]. The initial and final time spent by the animal in both chambers is recorded. Tendency of a subject to spend more time in the nonsocial chamber is a feature of autistic subjects [52]. Animals with autistic-like behavior are reported to spend more time in the chamber that lack social stimulus while healthy animals tend to spend more time in the chamber that has social stimulus [51,53].

3.2. Tests for Social Communication Behaviors

3.2.1. Ultrasonic Vocalization

USVs have emerged in rodents as valuable behavioral biomarkers in neurodevelopmental research. Rodents emit complex USVs ranging from 20 to 110 kHz, fulfilling various social and emotional functions and reflecting their communicative abilities [54]. These ultrasonic vocalizations in mice, observed across diverse social contexts including mother-pup interactions, juvenile play, adult courtship, same-sex social investigations, and territorial encounters, exhibit significant developmental changes throughout the lifespan, mirroring certain aspects of human social and communicative development. This development is associated with thermoregulation and neural maturation and is also influenced by genetic and environmental factors [55,56].

Recording and Analysis of USVs

USVs recordings require specialized equipment and controlled environmental conditions. Standard recording protocols employ ultrasonic microphones with a 10-180 kHz frequency response, positioned 10-30 cm above the test arena within acoustically isolated chambers to ensure precise sound capture and minimize noise interference. Environmental temperatures are typically maintained between 21-34°C, depending on the protocol, with colder temperatures sometimes used for higher stress induction [18,57]. Circadian rhythm disruptions frequently appear in many neurodevelopmental disorders and testing for different circadian behavioral changes is recommended [58].

USV analysis offers one of the few reliable means to assess communication behaviors during early postnatal life [59]. The application of machine learning has further improved the accuracy and throughput of USV analysis, enabling a more comprehensive characterization of vocal repertoires [54]. Advancements in automated USV analysis systems, such as those using machine learning algorithms, have significantly improved the accuracy and efficiency of identifying and classifying call parameters like frequency, duration, and amplitude. These parameters, may serve as reliable indicators of social and emotional states in rodents, providing a quantifiable measure of their communicative abilities [60,61].

USVs in Behavioral Testing Paradigms

Isolation Tests: Isolation tests are among the most prevalent methods for assessing social and communication deficits. These tests involve separating pups aged 3-14 days from their mothers and littermates typically for a duration of 3 to 10 minutes to elicit distress calls. The rate of calling as a function of age in healthy pups is typically described as a shallow inverted U-shaped pattern, appeared around day 3, peaking around day 5 and disappearing around day 12, while rat pups show a longer span of 15-17 days. Parameters such as call rate, duration, and frequency are measured, revealing significant differences between ASD models and controls [18,62,63]. This method is sensitive to ASD-related phenotypes and provides insights into the neurodevelopmental impacts of genetic and environmental factors on early communication behaviors [62]. For example, a number of studies have shown a disruption of the typical U-shaped calling rate function in mice with NDD [64,65].

Social Interaction Tests: These tests evaluate social communication behaviors by quantifying USVs emitted during social encounters such as pairings of unfamiliar adult rodents, maternal - offspring interactions, or courtship behaviors. ASD models often exhibit reduced frequency and complexity of these calls, indicating impaired social interaction and distress communication [66]. USVs produced during these interactions serve as a proxy for social motivation, communication, and recognition. Alterations in the frequency, complexity, and structure of these calls in ASD models often indicate impaired social interactions and emotional processing. For example, male rodents typically emit complex USVs during courtship, while female rodents may respond with specific vocalization patterns. In ASD models, these courtship-related USVs are often reduced in frequency and complexity, reflecting deficits in social behaviors and communicative intent [67].

ASD model rodents, such as Shank3 mutant and Cntnap2 knockout mice, show significant reductions in the number of USVs during social interactions, correlating with decreased social engagement and motivation [68,69]. Additionally, ASD model rodents often exhibit simpler spectrographic features, such as fewer modulations in pitch or shorter durations, suggesting diminished communicative flexibility [68]. These alterations in USV patterns suggest impaired social motivation, communication, and emotional processing in ASD models.

USV Changes in ASD Pups Following Separation from Their Dam and Nest

Rodent pups emit USVs when isolated from their mothers, serving as critical signals for maternal retrieval and social bonding. Alterations in the quantity, frequency, duration, and timing of these vocalizations have been regulated by genetic and environmental factors resulted in impaired communication development associated with ASD [62–64,67]. For instance, studies involving Cntnap2, as well as SHANK3 knockout rat models demonstrated that both genetic mutations and environmental challenges, such as maternal immune activation, can exacerbate changes in USV characteristics, including pitch and call type transitions [70,71]. BTBR T+tf/J mice, an inbred strain used as a model for idiopathic ASD, exhibit significantly fewer isolation-induced ultrasonic vocalizations (USVs) early postnatal stages and demonstrate unusual spectral and temporal features in their calls, suggesting impaired social communication development [64]. In addition, the atypical USV sequences, comprising both simple and complex call types, were observed in newborn mice with a Tbx1-heterozygous genetic model associated with ASD. These abnormal vocalizations disrupted effective social communication between pups and their mothers, highlighting the potential of USV analysis as a tool for early detection of ASD-like behaviors in animal models [72]. These alterations in USV profiles during early postnatal development may serve as valuable behavioral biomarkers for ASD-like phenotypes and provide insight into the neurobiological mechanisms underlying social communication deficits.

USV has been widely used to demonstrate behavioral deviations as well as altered emotional states. For example, adult rats have shown a modulation of increasing USV frequency from 22-kHz to 50 kHz with stressed negative vs positive emotions respectively, with distinct acoustic parameters characterizing the different states [73]. A similar distinction of 40 kHz vs 60 kHz frequency classes has been documented in juvenile or adult mice, with lower frequencies associated with aversive conditions and higher frequencies associated with a large variability in call characteristics and subtypes of calls [74]. Depression like behavior has been associated with USV changes, including a reduction in the number of typical stress emitted calls, namely flat and one-frequency step-up calls, and an increase in step-down type calls, associated with lower levels of stress and anxiety [75]. Evidence for a reduction in USV frequencies has been shown in rat models of schizophrenia [76,77], and acoustic impairments have been also demonstrated in mouse models [78,79].

Common Pitfalls in USV

USV studies in rodents are subject to methodological challenges that can significantly impact data interpretation. Subtle changes in the experimental setup, including ambient temperature, humidity and inconsistent recording environments, as well as variations in social context, and

experimenter-induced stress, or lack of experience may result in significant altered vocal patterns [64,66,67]. Technical limitations in recording equipment may also play a critical role in USV studies. Inadequate microphone sensitivity or improper positioning can lead to missed or distorted vocalizations, thereby skewing data. Recent advancements suggest that combining data from multiple high-quality microphones can enhance the accuracy of USV detection and localization [80]. Data analysis and classification may present another set of challenges. Traditional methods often rely on manual identification and classification of calls, which are time-consuming and subject to human error [80].

Species and strain variability: Mice and rats emit USVs that may express their unique arousal and emotional states. Rodent species and even strains within a species can exhibit variability in USV characteristics and findings in one strain may not be generalizable to others. Age and developmental stage considerations can also influence USV emission. Infant and adult rats, for example, emit USVs in different contexts, and even when emitted in response to the same stimulus, these vocalizations can have different characteristics [81]. Therefore, age-matched controls and longitudinal studies are essential to accurately assess developmental changes in USV patterns.

When designing studies to test the behavior of rodent pups using USVs after maternal separation, it is crucial to consider the duration of separation, as inconsistent results have been reported due to variations in separation protocols. Short periods of separation may not constitute serious deprivation, while longer separations can induce stress in pups and potentially lead to more pronounced behavioral changes [82]. However, excessively long separations or complete isolation from both mother and littermates may cause severe neural stress, potentially confounding the results and making it difficult to distinguish between ASD-like behaviors and general stress responses [82,83]. Finally, the lack of standardized protocols across research laboratories compounds these issues, making cross-study comparisons challenging and potentially leading to spurious conclusions about rodent behavioral and communicative processes.

To meet these challenges, it is advised to employ advanced analytical tools and machine learning modalities when possible, and to provide thorough training for experimenters to reduce variability and enhance reproducibility.

3.2.2. Scent Marking Test

The scent marking test aims to investigate communicability in rodents. Naturally, rodents communicate via olfactory signals such as scent markings and pheromones [48]. During the experiment, the time spent by the test animal sniffing the scent marking and the number of times it deposits urine near the scent marks is a measure of willingness to communicate. Control and wild type rodents demonstrate significant communicativeness compared to models of ASD-like phenotype [84–86]. The paradigm can be modified for habituation and dishabituation, whereby the animal is habituated with a social odor (pheromone) in the first stage and a nonsocial odor (rose or vanilla) in the second stage. In the final stage, the animal is presented with both odors (social and nonsocial) and the time it spent sniffing either of the odors is measured. Socially defective models are reported to show less/no preference for the social odor while normal/wildtype rodents preferred the social odor [51].

3.2.3. Pitfalls to Social Behavioral Tests

The 3-chamber test has no major setback in the design, principle and principal goal of its application. However, considering the significant rise in conflicting results obtained from social behavioral tests among researchers using this paradigm, there is a need to standardize the protocol and the apparatus to facilitate reproducibility and reliability of outcome. The partition test has challenges relating to apparatus and interpretation of outcome. Test animal could spend more time close to the partition without intention of interacting with the conspecific but rather because the partition spans the entire width of the arena relative to the small size of the rodent. In order to overcome this challenge number of intentional approaches towards the partition should be a

parameter of the test. The assumption that olfaction is the primary means of communication is accurate. However, findings have shown that opacity of the partition can reduce the socialization index of rodents by 30% [41]. One of the major pitfalls of dyadic social interaction is that the presence of an observer in the room can influence the test performance. The complex nature of the parameters observed in the test session make automation a near impossibility leaving the test in the hand of human observer prone to bias/subjective scoring and errors. It is possible to introduce videos into the recording room to eliminate the presence of an observer. Although the introduction of video recording would solve the problem of observer presence, it cannot prevent errors from manual judgment of the task. The latter is true on realization that rodent encounters are problematic *especially in all-males as fights, aggression and anxiety often ensue to complicate the definition of test parameters* [46]. While there are several methods to accomplish the objectives of scent marking test, the olfactory habituation/dishabituation (OHDH) is demonstrated to be the best paradigm of choice because of its ability to discriminate social deficit from olfactory deficit [44]. In the case of other approaches of the test, it is recommended that pup nest - seeking or olfactory discrimination test should be performed early in development to dismiss doubts of olfactory impairment that might arise in the future. Recommendations for a scoring method is plausible, considering that social behavioral deficit is a shared attribute among NDDs including ASD [5,29,30], ADHD [29,32], OCD [36] and SZ [87–89].

In studying social behaviors, it is crucial to employ multiple corroborating tests to effectively identify social impairments and ensure a comprehensive evaluation across various behavioral domains, including social, motor, cognitive, and anxiety-like behaviors. This multifaceted approach is necessary to reveal unconfounded NDDs such as ASD-like phenotype, as no single task can adequately measure social behavior or communication. Relying on too few tests may lead to artifactual results, thereby masking the significance of potential confounds in behavioral assessments.

3.3. Test Paradigms for Restricted Repetitive Behavior

3.3.1. Marble Burying

To examine rodents for the presence of obsessive/compulsive behavior, typical in autism, ADHD, anxiety and OCD, marble burying test is employed frequently [47,90]. Repetitive behaviors in rodents often manifest as excessive digging behavior when presented with glass balls in a novel arena [91]. Repetitive and compulsive-like digging behavior is measured by counting the number of balls a test animal buries during a test session [92,93]. The experimental set up consists typically of 20 glass/marble balls, a standard cage with clean bedding materials. Clean cages are filled up to 4-5 cm with bedding material, the balls are gently placed in 4 columns by 5 rows on top of the bedding. The test rodent is then introduced to the arena from one corner to prevent balls scattering. The test duration majorly is 30 minutes during which the rodent is left alone without interference to interact with the balls. At the end of test, the number of marbles that the subject buried by 2/3 of their diameter is used to measure the repetitive digging behavior [37]. NDDs models, especially autistic-like models, have a high tendency to bury more balls than wildtype and control animals [94,95].

3.3.2. Barnes Reversal

Reversal of Barnes maze paradigm is used to examine behavioral inflexibility and spatial memory in rodent models of NDDs such as ASD [5,44]. The apparatus consists of a circular platform (92 cm diameter) with 20 holes of 5 cm diameter each equally spaced at the perimeter. One of the holes leads to a dark escape box that protects the test subject from brightly illuminated test room. In experiments assessing repetitive behavior, the protocol consisted of five phases (1) adaptation (2) forward acquisition training (3) forward test trials (4) reversal training, and (5) reversal test trials. During adaptation, test subject is placed in the middle of the platform and covered for 10 seconds with a dark box. This is followed by the 2nd phase during which the subject is uncovered and trained to learn the location of the hole housing the escape box within 3 minutes. After learning the position of the escape hole, the 3rd phase is when the subject is introduced and allowed 90 seconds to find the

location of the hole during which its performance is scored. Investigators can measure the latency to find the correct hole and the number of incorrect holes explored (hole poking error, and the length of the exploratory path [39,96,97]. In the reversal phases, the experiment is basically repeated in later days with the position of the escape hole changed to 180 degrees. The training and test are repeated with the escape position reversed and once again the scores are measured [98,99]. The maze could be divided into 4 quadrants so that the subject duration in the escape hole quadrant is measured as well [100].

3.3.3. Repetitive Stereotype Behaviors

Repetitive behaviors in rodents can be observed in normal activities like self-grooming and bedding [98]. In rodents exhibiting ASD or OCD like behavior, these activities continue for unusually longer time than normal pattern self-grooming [101,102]. Stereotypic behaviors like backflipping, circling, and jumping are performed more frequently in animals exhibiting autistic like behavior compared to healthy ones [45]. Following habituation, a subject is observed and scored for cumulative time spent grooming, digging or chewing. Time spent exhibiting a particular action between models of NDDs and normal subjects are compared as index of impaired repetitive behavior [102].

Pitfalls to Restricted Repetitive Behavioral Test

The marble burying test has some challenges facing the validity of results obtained from the studies mainly due to the multipurpose application of the test outcome. Sometimes excessive digging is related to anxiety disorder and in other cases is judged as OCD symptoms [103]. Most recently, the test has gained popularity and acceptance among autism researchers as an important tool for studying repetitiveness in rodent models of ASD-like phenotypes [5,37,104,105].

Questions have been continually raised regarding the validity and reliability of marble burying test outcomes. One of the few problems is lack of standardized arena dimensions which stem from the fact that the test is performed classically using home cage. Since different researchers use varying cage sizes, there is, actually, arena size effect which affects the number of balls and the proximity of balls to each other. Because of this, there are balls that are often covered by bedding material from digging adjacent balls. The nature of bedding material also matters as saw dust and wood shaving spread more further from digging than other types of bedding materials. Some investigators allow the task to proceed with the cage open (with some animals leaving the cage and being returned) while others cover the cage with a lid to overcome the eventual escape of mice that occurs in much of the cases. The Barnes test is primarily used to assess spatial memory in rodent models of cognitive impairment. However, the introduction of the reversal aspect empowers the paradigm to simultaneously assess both spatial memory and repetitiveness in rodent models of NDDs and other psychiatric diseases [106]. The Barnes maze, although lacking spatial continuity, has the advantage of a dry platform which removes the stress of water and force swimming that are the major pitfalls for Morris water maze (MWM) [97,99]. Impaired memory has been reported as one of the comorbidities of ASD and other neurodevelopmental and neuropsychiatric disorders [44,96]. Therefore, the Barnes Maze paradigm is considered very advantageous, not only for the less stressful environment it offers but also for the benefit of studying memory together with repetitiveness. The presence of an investigator might be a conflicting source of external influence on the animal's behavioral task performance. It is important for the observer/examiner to minimize movement and meticulously ensure all test subjects, and experimental groups received equal measures of all manipulations in each test performance.

3.4. Tests for Cognitive Impairment

3.4.1. Morris Water Maze (MWM)

The MWM is a test paradigm for assessing spatial learning and memory and reversal learning in rodent models of NDDs [107]. With the aid of spatial cues, subjects learn the location of a platform hidden underwater as the escape route from a circular pool. The apparatus is a round water pool of about 100 cm in diameter and water of temperature 22°C. The hidden escape platform is placed 1-2 cm below the water in one quadrant of the circular pool, and the water is colored to conceal the platform. Rats are trained for five consecutive days in the apparatus to learn the location of the platform by swimming using distinct visual cues placed around the maze [108]. During the training days, mice are introduced into the pool four times per session via four different locations and the speed and latency to locate the platform is recorded. Each trial time elapses when the subject escapes to the hidden platform or is placed on it by the experimenter if it could not locate it. On the test day after training days, the test is performed again, this time with the platform removed [109]. The time the subject spent in each quadrant, and the number of times it crosses the original location of the platform is recorded. Preference ratio can be computed as; time spent in platform quadrant / time spent in the other three quadrants [34,96,108].

3.4.2. Y-Maze Spontaneous Alternation

This test is employed in rodents to measure hippocampal function in the context of short-term memory (working memory) [96]. The test arena is a maze that consists of three identical arms (forming Y shape) of 40 cm length, and 10 cm width and height each. During experiment, animals are placed in one of the arms from which it is allowed 5 minutes or more to sequentially explore other arms freely. Test animals with intact working memory will normally alternate the arms after each visit by moving unto a new arm and not going back to a previously visited arm. An animal with impaired cognition tends to show reduced spontaneous arm alternation that accompanies increased arm revisit which indicates impairment in memory events [34]. Percentage alternation is calculated as ratio of the total number of alternations (triad containing entry into all three arms) to the total number of entries minus 2 multiplied by 100 [110]. Deficits in spontaneous alternations are characteristic of the ADHD and schizophrenia phenotypes and could be ASD-like comorbidity features.

3.4.3. Novel Object Recognition (NOR)

The NOR is also used to assess intellectual difficulties that may be present in most NDDs and confounds ASD [44]. The test which measures memory is based on rodent's innate tendency to explore novel objects than a familiar one [96,111]. (The apparatus is an open box of about 50 × 50 × 40 cm. The test proceeds by first familiarizing the subject with the arena at a given time (5 minutes). This is followed by introducing two identical objects (including size, color and shape) and the test subject is allowed to familiarize itself with them [112]. Finally, one of the familiar objects is replaced with a novel object (only size should be maintained) and the subject is allowed to explore the familiar versus the unfamiliar object. Placing objects in the same position is important to avoid localization-related differences. The time spent exploring novel and familiar objects are recorded and discrimination index computed as; time with a novel object - time with a familiar object) / (time with a novel object + time with a familiar object [113,114]

3.4.4. Fear Conditioning

The test is used to assess fear processing, and to associate an aversive foot shock with contextual or sound cues in rodent models of autistic-like phenotype. NDD models are reported to display elevated fear memory and generalization for sound and contextual cues [115]. The protocol begins with a habituation phase when subjects are allowed to acclimate to the conditioning chamber without stimuli. After habituation, the conditioning phase involves presenting a sound cue for 30 seconds followed immediately by an aversive shock, repeated several times with inter-trial intervals and then returned to home cage. After conditioning phase, subjects undergo a contextual fear test, where they are placed back in the chamber with the sound cues but without the aversive shock [15]. Since rodent express fear as freezing behavior, time spent when in freezing state is taken as index of fear. Data analysis focuses on quantifying freezing behavior during both tests, with appropriate statistical methods employed to evaluate differences between conditioned and control groups [116,117].

Pitfalls to Tests for Cognitive Impairments

Y-maze test is a valuable tool for assessing spatial memory, anxiety and repetitiveness in rodents, yet it is considered not to be as definitive as the Barnes maze and other more specialized learning/memory tests [98]. Pitfalls in Y-maze could include variability in individual behavior as some animals may exhibit inherent differences in exploration tendencies or anxiety levels. The novelty of the environment and maze design can influence behavior, potentially confounding results. It is important to allow familiarization time of 1-2 minutes before recording subject performance. The reliance on spontaneous alternation may also affect the ability to discern between memory and motivational factors [118]. There are inconsistencies between laboratories in terms of lighting conditions, maze size and test duration that needed to be addressed to lessen the gap in results data. Therefore, careful consideration of these variables is essential for robust conclusions in studies utilizing the Y-maze test.

There are a few challenges in relation to MWM despite it being the most robust and reliable test for assessing spatial learning and memory. The aquatic stress involved can impact rodent performance because of differences in anxiety and swimming ability. This presence of water could lead to extraneous behaviors quite unrelated to spatial learning, particularly in mice [96,98]. Too much focus on escape time in addition to individual subjects' escape strategy could mask subtle memory impairments. Lack of report/regulation on water temperature and maze size is commonly practiced in many studies regardless that these are culprit to task performance and result variability. Although considered the cornerstone of the field, careful consideration of these factors is crucial for valid conclusions in MWM studies.

In NOR, Variability in objects explored can affect replicability, as some rats may show little interest in the objects, necessitating strict inclusion/exclusion criteria based on performance levels. The transient nature of novelty means that attention to new objects may diminish quickly, which impacts the reliability and validity of results. It is important therefore to optimize the test duration as many studies fail to control the time spent with objects during familiarization, potentially skewing retention data and leading to inaccurate conclusions. It is also important to acclimatize the subject with the arena before any object introduction.

Fear conditioning test measures for associative and working memory. Contextual and cued conditioning components of the paradigm mainly assess specific associative learning, potentially overlooking other cognitive responses. The specific nature of cues can limit wide applicability of results. Moreover, the fear conditioning test has a major pitfall in the reliance on foot-shock as a stressor capable of inducing significant anxiety and distress in the subjects. This raises ethical concerns regarding animal welfare, and therefore compliance with animal welfare guidelines and minimizing distress must be considered as paramount. The level of individual subject's response to stress differs and may introduce freezing behavior that are not components of memory.

3.5. Tests for Sensory Stimuli

3.5.1. Prepulse Inhibition (PPI)

Animal models of psychiatric disorders exhibit central deficits in sensorimotor gating, (i.e., abnormal PPI) of the robust acoustic startle response (ASR), which can be reliably measured [88,111,119]. The apparatus consists of two soundproof test chambers of about 40 x 40 x 60 cm dimensions, restrainer in the chambers, loudspeakers to generate acoustic stimuli and a sensor to measure startle. The test is made up of a habituation day during which test subject is acclimatized to the test environment. Acclimatization includes exposing subjects to a steady background noise of 70 dB (decibel) and multiple startle stimuli of 120 dB to test hearing and motor functions. On the test day the subject is exposed to the arena and recorded over several exposure to a weaker acoustic stimulus (prepulse) followed by an intense startle stimulus (pulse). A sudden puff of air can be used as a tactile stimulus to elicit startle response which can then be quantified [119]. By measuring startle amplitude at the beginning and end of a session, the percent prepulse inhibition can be calculated using the formula $\%PPI = 100 \times ([\text{startle-only units} - (\text{pre-pulse} + \text{startle units})] / (\text{startle-only units}))$.

3.5.2. Nociception

This paradigm is used to assess rodent reactivity to sensory stimuli. To achieve this purpose, the hot plate and tail flick tests are frequently employed. The subject is basically placed on a plate that is heated and kept at a constant temperature of about 55 °C. The latency to the first reaction (licking of paw, jumping) of the animal is recorded. Similarly, in the tail flick test the subject tail is placed in hot water or beam of light and the latency to flick the tail is recorded [119,120]. Rat models of postnatally induced ASD display increased pain insensitivity with delayed first attempt to lick the paws on a hot plate [120].

3.6. Tests for Anxiety-Like Behavior

3.6.1. Elevated Plus Maze (EPM)

Elevated plus maze (EPM) and light-dark box tests are suitable and widely used test for anxiety [16,120]. These tests explore the approach-avoidance conflict between mice innate curiosity to explore novelty and mice preference to be in closed/dark instead of opened/illuminated places. The apparatus is a plus-shaped maze whereby each arm measures about 30 cm in length, 8 cm in width. Two opposing arms have 17 cm high opaque walls around them except the central entrance (closed arms); the other two opposing arms are not walled (open arms). This arrangement creates a central platform to which animals are introduced for observation. The maze is pivoted at a height of 70-100 cm above the floor to discourage test subject from jumping off the platform. During the experiment subject are placed on the central with its face towards the open arm [121]. The subject is then given time to explore the arena for a given time according to the experimental design [108,122].

3.6.2. Dark/Light Box Tests

The dark-light apparatus is made up of a white painted and brightly illuminated open box, that is connected to another black painted non-illuminated closed box. The boxes are rectangular of varying dimensions from lab to lab. During behavioral observation, subjects are placed gently in the middle of the white box with the faces directed opposite to the connecting opening [15,123]. The frequency of entries and the total time spent in the open/lighted, the closed/dark region, and the central zone is recorded and analyzed [124]. The increased time spent in open arms of EPM or in the illuminated part of light-dark box is an indication of lower degree of "anxiety" in the animal models of autistic-like behavior [134].

3.6.3. Open Field Locomotory (OF) Test

This test is widely employed for the assessment of general locomotion, exploration and anxiety behavior. The paradigm consists of a square arena that is divided into 9 equal parts [107,121]. For the purpose of examination, a subject is introduced to the center of arena and allowed 5 to 10 minutes to explore the open field, and its activities are observed and recorded. Locomotory/ exploratory behaviors like horizontal and vertical activity, movement and rest time, and total distance travelled are measured [88,124]. The latency, duration, and the number of visits spent in the center of the arena or on the wall, are recorded for computing the anxiety level [15]. In addition, repetitive & stereotyped behaviors of subject can be determined by observing the self-grooming, spinning, jumping and rearing behavior [122]. Rodent models of NDDs especially ASD-like and ADHD-like models are reported to spend more time grooming in abnormal fashion, spend more time near the walls than the center of the arena [33,94,101]

3.6.4. Pitfalls to Anxiety Tests

Studies that involve behavioral tests for anxiety-like behaviors including EPM, Light- Dark Box, and OF tests, are advised to measure and report that lighting conditions are optimal and consistent to ensure that animal responses align with expected outcomes. Consistent reporting of species, age and sex is crucial for understanding these behaviors. Meta data analysis showed that studies have reported duration as low as 4 minutes to as high as 30 minutes pointing to the lack of standardized optimal latency and duration in the application of these tests [125]. Most studies use test durations of 5 to 10 minutes; however, novelty in a new environment can induce neophobia, skewing results toward increased anxiety due to insufficient habituation. Conversely, exceeding 10 minutes may reduce anxiety responses as animals acclimate. Therefore, a 10-minute duration is recommended, with the first 5 minutes assessing anxiogenic responses and the last 5 minutes evaluating anxiolytic responses.

In the Elevated Plus Maze, lighting and noise can significantly influence performance and alter anxiety responses, as the test primarily measures approach-avoidance behavior, which may not fully capture the complexity of anxiety disorders. Similarly, in the Light and Dark Box, rodents' preference for dark environments can confound interpretations, and environmental factors like light intensity and arena size may also affect results. The Open Field Test may not distinguish between anxiety and general activity levels, as increased movement can arise from curiosity or fear. Thus, careful consideration of these factors is essential for valid conclusions in anxiety-related studies.

3.7. Tests for Motor Coordination and Balance

3.7.1. Rotarod Paradigm

The rotarod test is a conventional assessment tool for motor coordination in rodent-based experimental studies on NDDs and other brain disorders that involve impairment in motor coordination [44,96]. Using a rotating rod paradigm, motor coordination and balancing as well as motor skill learning can be tested by recording the amount of time a subject remains moving on the rod while an improved performance across repeated test sessions is a marker of motor learning ability [126]. The rotarod apparatus consists of a rotating rod that accelerates in predetermined time intervals. The floor is padded to prevent injury from falling offs because the rod lies horizontally at height that is high enough to scare subjects from jumping [127]. The latency to fall over varying speeds are recorded as indications of the motor coordination function and similar judgement is applicable to beam walking (static rod) test of motor function in rodents [128].

3.7.2. The Beam Walking Test

This is a widely used alternative paradigm to rotarod suggestive of better sensitivity than the rotarod test in detecting deficits in motor coordination because in addition to latency, the test uses quantitative values such as walking distance and number of slips to accomplish the task on the

narrow beam [129,130]. The beam walk apparatus is a wooden beam of 80 cm of length and 1 cm of width at a height of 60 cm. At one end of the bar is attached a dark escape box at 10 cm from the finish line while a line 10 on the free opposing end marks the starting point. During training and test days, test subject is placed on the elevated beam from the free starting end and the latency to cross the beam and number of foot slips are recorded as measures of incoordination in rodents of NDDs [131].

The beam walking paradigm is not popular in terms of application compared to the rotarod. However, it appears to have the additional advantage to assess balancing and sensorimotor gating. It also has low-cost advantage and a simple paradigm

To further probe the deficit found from rotarod and beam walking tests, other tests of coordination like footprint analysis of locomotor gait, mesh climbing, swimming, pole climbing, and the staircase test can be employed

3.7.3. Pitfalls to Test for Coordination and Balance

The rotarod test, even though widely used for assessing motor performance in mice and rats, has several notable limitations that can complicate the reproducibility and interpretation of results. One of such issue is that, while some animals may cling to the beam and rotate alongside the rod instead of properly balancing on the rotating rod, others may fall immediately upon being placed on it. Additionally, the apparatus may not accelerate sufficiently and uniformly to accurately detect motor incoordination effectively, and if the height is not optimally raised animals can be encouraged to jump off the rod. Variability in test results often arising also from different test paradigms and apparatus designs (including lighting condition, diameter and texture of rod, and width of the compartments), which can yield inconsistent outcomes under similar conditions. Furthermore, certain strains of mice are known to perform poorly on the rotarod, raising questions about the validity of conclusions drawn from such tests. Factors like stress, learning effects from repeated trials, and test environment can further influence results. It is crucial to standardize training protocols and consider the mentioned factors to ensure reliable assessments of motor coordination.

As animals often jump off during task performance, it is important to ensure that the beam is elevated high enough. Lack of motivation can be a confounding factor if the subjects are subjected to multiple training that can translate to uncooperativeness. The surface of the beam should have a uniform smoothness without any textures or protrusions that subjects might grip on. Animals from different test groups should be of the same age, weight, strain and weight. Weight difference might be an important factor in this test paradigm; considering that older and heavier animals tend to be slow and likely to have more slips on a beam of equal width. Finally, for the purpose of assessing balancing in rodents, it is important to consider the width of the beam because a too wide beam may not necessary or effectively assess balancing.

4. Discussion

We have presented a variety of methods assessing different behavioral traits in rodents that are used to measure behaviors which characterize neurodevelopmental and psychiatric disorders. – NDD). For each set of behavioral tests, we described the possible pitfalls stressing the likelihood for errors that might affect the results. While deviations in many of these behaviors are found in ASD, most of them are not specific to ASD like behavior as they may also characterize other neuropsychiatric disorders. Hence, in many NDD, the existence of a specific combination of such deviations points to the diagnosis of a specific disorder. However, for a precise diagnosis, a minimal number of typical behaviors exhibiting a significant severity must be present.

DSM 5 and other diagnostic measures require to grade the severity of the observed behavioral deviations for the proper diagnosis. Indeed, the behavioral assessment tools used for diagnosis generally also grade the severity of the deviations. This is true, for example, with the common diagnostic tools for ASD such as the Child Autism Rating Scale (CARS), Autism Diagnostic Observation Schedule (ADOS), Autism Diagnostic Interview-Revised (ADI-R), and other diagnostic tools [1].

The different tests used for the assessment of behavioral changes in animals, especially in rodents, are generally adequate to assess all specific and typical autistic-like behaviors. However, studies assessing only a few behavioral tests for the demonstration of a specific NDD (i.e., only tests for social interaction and communication or repetitive behaviors and anxiety to demonstrate ASD-like behavior) may be insufficient. Such studies can be used, however, to define specific traits of autistic-like behavior: i.e., communication difficulties, repetitive behavior, restricted interests, abnormal response to sensory stimuli and others, but not the complete diagnostic set. Hence, an accepted scoring system or at least, definition of the severity, similar to that in humans, seems to be important for all animal models that mimic human neurobehavioral and neuropsychiatric diseases.

In the last years there are, indeed, several attempts to standardize behavioral data and provide relative severity to behavioral tests and gender differences.

4.1. Integrated Behavioral z-Scoring

In the research of NDD, z-score can be used to standardize complex pathophysiology into a single paradigm, similar to diagnosing mental health conditions in humans. z-scores convert diverse measurements to a standard scale by subtracting the population average from an individual raw value and dividing by the population standard deviation, allowing researchers to compare and interpret data across different studies and variables [132,133]. Thus, Guilloux et al. [132] used this technique to normalize multiple anxiety- and depressive-like behavioral tests of paradigms, into a single emotionality z-score index. The authors emphasize the significance of z-normalization approach as an analysis that increased the sensitivity and reliability of behavioral phenotyping in mice [132].

Kraeuter et al., [133] described the methodology of integrated behavioral z-scoring in rodent studies. Integrated behavioral z-score can be applied to the psychiatric or neurodevelopmental disorders, such as ASD, and even to combine it with biochemical and molecular changes z-scores, forming a comprehensive disease specific integrated z score.

Labots et al. conducted a study using the modified Hole Board (mHB) test to evaluate behavioral characteristics in three inbred mouse strains: A/J, BALB/cJ, and C57BL/6J [134]. To analyze the data, they employed z-score calculations using the pooled data from all groups as a reference. This approach was needed as the experimental design lacked a common control group due to the comparison of both sex and strain differences. Labots et al. noted that traditional z-normalization methods would be impossible if the control group had a standard deviation of zero, highlighting the advantage of their pooled data approach [134].

El-Kordi et al. utilized z-score normalization to analyze ASD-like behavior in *Nlgn4* null mutants, with higher z-scores indicating greater symptom severity [135]. For males, the autism severity score encompassed seven behavioral categories spanning the three core symptoms: qualitative impairments in social interaction (social approach behavior, nest building, and aggression), communication deficits (ultrasonic vocalizations during male-female interactions), and restricted, repetitive, and stereotyped behaviors (marble burying and circling behavior) [135]. Similar behavioral categories were identified as relevant for females [136]. Applying the autism severity score demonstrated near-perfect genotype classification for individual male mice (almost 100% accuracy), while accuracy for female mice was slightly lower, at approximately 80% for correct group assignments. These findings highlight the utility of z-score normalization as a robust method for quantifying and distinguishing ASD-like behaviors across genotypes.

In our previous studies on ASD-like behavioral changes induced by early postnatal administration of VPA, we developed a composite scoring system to classify these behavioral alterations [96]. We integrated z-scores from self-grooming frequency in the open field test, incorrect turning percentages in the T-maze, and social novelty preference in the three-chamber social interaction test as three key behavioral measures. Our findings showed that the autism composite score was significantly higher in VPA-treated male and female mice compared to controls and was normalized by the co-administration of SAME. Despite these findings, the study did not address the

severity of individual behavioral changes, which could be assessed by quantifying the degree of deviation from control values for each animal.

In clinical practice it is advised to use a 4- scales scoring system in the definition of the severity of a NDD: normal, slightly abnormal, moderately abnormal, severely abnormal [1,5]. The possible scores for each behavioral characteristic are from 0-3. Zero being normal behavior and 3 - severely abnormal. As different NDD are characterized by abnormalities in several specific behavioral traits, a composite z-score for all behavioral characteristics studied is generally inadequate. Hence, it is expected that each one of the typical behaviors tested will have its own severity score, and then calculate the composite score, as used in clinical medicine. Moreover, it is expected that for each NDD, most, if not all of the behaviors expected to be abnormal, will be tested. For example, to define an appropriate animal model for autistic-like behavior, it is imperative to assess sufficient behavioral traits that characterize ASD like behavior; at least two of the three specific social characteristics (social behaviors, social interaction and communication), and two of the following: restricted and repetitive behaviors, cognitive impairment, anxiety like behavior and motor coordination. In addition, there is a need for individual scoring of each of the behaviors tested.

As there are differences in the clinical presentations of some NDD between males and females, it seems important to carry out behavioral studies on both sexes.

5. Conclusions

We have described a variety of behavioral studies in rodents commonly used to describe animal models of human NDD. We pointed out the possibilities of serious pitfalls in the performance and interpretation of these studies. Investigators should therefore be aware of these possible drawbacks and consider them carefully before carrying out neurobehavioral studies, as they should be reflected in the assessments of the results. We also pointed out the need to use of sufficient and aimed behavioral tests to define a model of a specific NDD. Due to gender differences in many human NDD, it is appropriate to study males and females. To mimic the clinical definitions, it is important to assess the severity of each specific behavioral change.

It is recommended that behavioral experts decide on these salient factors which may affect the progress of preclinical findings. Consensus can be reached on the various issues raised by us, reducing the possible difficulties interfering with the translation of animal behavioral deviations to the human disorders.

Author Contributions: Conceptualization, A.O.; methodology, A.O., B.E.; writing—original draft preparation B.E.; M.B.; D.S., and A.O.; writing—review and editing, A.O. and M.B.; supervision, A.O.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflicts of interest.

Abbreviations

The following abbreviations are used in this manuscript:

NDD	Neurodevelopmental disorder
OCD	Obsessive compulsive disorder
ADHD	Attention deficit hyperactive disorder
ASD	Autism spectrum disorder
TS	Tourette syndrome
ZF	Schizophrenia
PND	Postnatal day
VPA	Valproic acid

CNS	Central nervous system
USV	Ultrasonic vocalization
OHDH	Olfactory habituation and dishabituation
MWM	Morris water maze
NOR	Novel object recognition
PPI	Prepulse inhibition
ASR	Acoustic startle response
DB	Decibel
EPM	Elevated plus maze
OF	Open field
CARS	Child Autism Rating Scale
ADOS	Autism Diagnostic Observation Schedule
ADI-R	Autism Diagnostic Interview-Revised
mHB	Modified hole board

References

1. American Psychiatric Association, D.; Association, A.P. *Diagnostic and statistical manual of mental disorders: DSM-5*; American psychiatric association Washington, DC: 2013; Volume 5.
2. Sonzogni, M.; ; Wallaard, I.; ; Santos, S.S.; ; Kingma, J.; ; Du Mee, D.; ; et al. A behavioral test battery for mouse models of Angelman syndrome: a powerful tool for testing drugs and novel Ube3a mutants. **2018**, *9*, 47, doi:10.1186/S13229-018-0231-7.
3. Eapen, V.; ; Mcpherson, S.; ; Karlov, L.; ; Nicholls, L.; ; Črnčec, R.; ; et al. Social communication deficits and restricted repetitive behavior symptoms in Tourette syndrome. **2019**, doi:10.2147/NDT.S210227.
4. Eyles, D.W. How do established developmental risk-factors for schizophrenia change the way the brain develops? *Translational Psychiatry* **2021**, *11*, 158, doi:10.1038/s41398-021-01273-2.
5. Ornoy, A.; Echefu, B.; Becker, M. Animal Models of Autistic-like Behavior in Rodents: A Scoping Review and Call for a Comprehensive Scoring System. *Int J Mol Sci* **2024**, *25*, doi:10.3390/ijms251910469.
6. Ornoy, A. The impact of intrauterine exposure versus postnatal environment in neurodevelopmental toxicity: long-term neurobehavioral studies in children at risk for developmental disorders. *Toxicology Letters* **2003**, *140-141*, 171-181, doi:https://doi.org/10.1016/S0378-4274(02)00505-2.
7. Jiménez, J.A.; Zylka, M.J. Controlling litter effects to enhance rigor and reproducibility with rodent models of neurodevelopmental disorders. *Journal of neurodevelopmental disorders* **2021**, *13*, 2, doi:10.1186/s11689-020-09353-y.
8. Zorrilla, E.P. Multiparous species present problems (and possibilities) to developmentalists. *Dev Psychobiol* **1997**, *30*, 141-150, doi:10.1002/(sici)1098-2302(199703)30:2<141::aid-dev5>3.0.co;2-q.
9. Festing, M.F.W. Design and Statistical Methods in Studies Using Animal Models of Development. **2006**, *47*, 5-14, doi:10.1093/ILAR.47.1.5.
10. Holson, R.R.; Pearce, B. Principles and pitfalls in the analysis of prenatal treatment effects in multiparous species. *Neurotoxicology and teratology* **1992**, *14*, 221-228, doi:10.1016/0892-0362(92)90020-b.
11. Dorsch, M.; Wittur, I.; Garrels, W. Efficiency of timed pregnancies in C57BL/6 and BALB/c mice by mating one male with up to four females. *Laboratory animals* **2020**, 23677219897687, doi:10.1177/0023677219897687.
12. Reddy, R.; Pillay, V.; Baijnath, S.; Singh, S.D.; Ramdin, S.; Naicker, T.; Govender, N. Mating success of timed pregnancies in Sprague Dawley rats: Considerations for execution. *Reproductive biology* **2022**, *22*, 100682, doi:10.1016/j.repbio.2022.100682.
13. Jafari, Z.; Faraji, J.; Mirza Agha, B.; Metz, G.A.S.; Kolb, B.E.; Mohajerani, M.H. The Adverse Effects of Auditory Stress on Mouse Uterus Receptivity and Behaviour. *Scientific reports* **2017**, *7*, 4720, doi:10.1038/s41598-017-04943-8.
14. Weinstock, M. Prenatal stressors in rodents: Effects on behavior. *Neurobiology of Stress* **2017**, *6*, 3-13, doi:https://doi.org/10.1016/j.ynstr.2016.08.004.
15. Eltokhi, A.; Kurpiers, B.; Pitzer, C. Behavioral tests assessing neuropsychiatric phenotypes in adolescent mice reveal strain- and sex-specific effects. *Scientific reports* **2020**, *10*, 11263, doi:10.1038/s41598-020-67758-0.

16. Tsao, C.H.; Flint, J.; Huang, G.J. Influence of diurnal phase on behavioral tests of sensorimotor performance, anxiety, learning and memory in mice. *Scientific reports* **2022**, *12*, 432, doi:10.1038/s41598-021-03155-5.
17. KSV, A.; Dissanayake, D.; Gunatilake, M.; Kuzhandai Velu, V.; Paranthaman, M. A short review on behavioural assessment methods in rodents. *Bioinformation* **2023**, *19*, 866-870, doi:10.6026/97320630019866.
18. Kahnau, P.; Mieske, P.; Wilzopolski, J.; Kalliokoski, O.; Mandillo, S.; Hölter, S.M.; Voikar, V.; Amfim, A.; Badurek, S.; Bartelik, A.; et al. A systematic review of the development and application of home cage monitoring in laboratory mice and rats. *BMC Biol* **2023**, *21*, 256, doi:10.1186/s12915-023-01751-7.
19. Saré, R.M.; Lemons, A.; Smith, C.B. Behavior Testing in Rodents: Highlighting Potential Confounds Affecting Variability and Reproducibility. *Brain Sci* **2021**, *11*, doi:10.3390/brainsci11040522.
20. Liu, Y.J.; Li, L.F.; Zhang, Y.H.; Guo, H.F.; Xia, M.; Zhang, M.W.; Jing, X.Y.; Zhang, J.H.; Zhang, J.X. Chronic Co-species Housing Mice and Rats Increased the Competitiveness of Male Mice. *Chemical senses* **2017**, *42*, 247-257, doi:10.1093/chemse/bjw164.
21. Yang, M.; Augustsson, H.; Markham, C.M.; Hubbard, D.T.; Webster, D.; Wall, P.M.; Blanchard, R.J.; Blanchard, D.C. The rat exposure test: a model of mouse defensive behaviors. *Physiology & behavior* **2004**, *81*, 465-473, doi:10.1016/j.physbeh.2004.02.010.
22. Chesler, E.J.; Wilson, S.G.; Lariviere, W.R.; Rodriguez-Zas, S.L.; Mogil, J.S. Influences of laboratory environment on behavior. *Nature neuroscience* **2002**, *5*, 1101-1102, doi:10.1038/nn1102-1101.
23. Ueno, H.; Takahashi, Y.; Murakami, S.; Wani, K.; Matsumoto, Y.; Okamoto, M.; Ishihara, T. Effect of simultaneous testing of two mice in the tail suspension test and forced swim test. *Scientific reports* **2022**, *12*, 9224, doi:10.1038/s41598-022-12986-9.
24. Saré, R.M.; Levine, M.; Hildreth, C.; Picchioni, D.; Smith, C.B. Chronic sleep restriction during development can lead to long-lasting behavioral effects. *Physiology & behavior* **2016**, *155*, 208-217, doi:10.1016/j.physbeh.2015.12.019.
25. Tanaka, T.; Ogata, A.; Inomata, A.; Nakae, D. Effects of different types of bedding materials on behavioral development in laboratory CD1 mice (*Mus musculus*). *Birth defects research. Part B, Developmental and reproductive toxicology* **2014**, *101*, 393-401, doi:10.1002/bdrb.21129.
26. Freymann, J.; Tsai, P.P.; Stelzer, H.D.; Mischke, R.; Hackbarth, H. Impact of bedding volume on physiological and behavioural parameters in laboratory mice. *Laboratory animals* **2017**, *51*, 601-612, doi:10.1177/0023677217694400.
27. Eltokhi, A.; Kurpiers, B.; Pitzer, C. Behavioral tests assessing neuropsychiatric phenotypes in adolescent mice reveal strain- and sex-specific effects. **2020**, *10*, 1-15, doi:10.1038/s41598-020-67758-0.
28. Peleh, T.; Eltokhi, A.; Pitzer, C. Longitudinal analysis of ultrasonic vocalizations in mice from infancy to adolescence: Insights into the vocal repertoire of three wild-type strains in two different social contexts. *PLoS One* **2019**, *14*, e0220238, doi:10.1371/journal.pone.0220238.
29. Carter Leno, V.; Chandler, S.; White, P.; Pickles, A.; Baird, G.; Hobson, C.; Smith, A.B.; Charman, T.; Rubia, K.; Simonoff, E. Testing the specificity of executive functioning impairments in adolescents with ADHD, ODD/CD and ASD. *European child & adolescent psychiatry* **2018**, *27*, 899-908, doi:10.1007/s00787-017-1089-5.
30. Ceruti, C.; Mingozzi, A.; Scionti, N.; Marzocchi, G.M. Comparing Executive Functions in Children and Adolescents with Autism and ADHD-A Systematic Review and Meta-Analysis. *Children (Basel, Switzerland)* **2024**, *11*, doi:10.3390/children11040473.
31. Rommelse, N.N.; Geurts, H.M.; Franke, B.; Buitelaar, J.K.; Hartman, C.A. A review on cognitive and brain endophenotypes that may be common in autism spectrum disorder and attention-deficit/hyperactivity disorder and facilitate the search for pleiotropic genes. *Neuroscience and biobehavioral reviews* **2011**, *35*, 1363-1396, doi:10.1016/j.neubiorev.2011.02.015.
32. Townes, P.; Liu, C.; Panesar, P.; Devoe, D.; Lee, S.Y.; Taylor, G.; Arnold, P.D.; Crosbie, J.; Schachar, R. Do ASD and ADHD Have Distinct Executive Function Deficits? A Systematic Review and Meta-Analysis of Direct Comparison Studies. *Journal of attention disorders* **2023**, *27*, 1571-1582, doi:10.1177/10870547231190494.
33. Chao, O.Y.; Yunger, R.; Yang, Y.M. Behavioral assessments of BTBR T+Itpr3tf/j mice by tests of object attention and elevated open platform: Implications for an animal model of psychiatric comorbidity in autism. *Behav Brain Res* **2018**, *347*, 140-147, doi:10.1016/j.bbr.2018.03.014.

34. Luhach, K.; Kulkarni, G.T.; Singh, V.P.; Sharma, B. Attenuation of neurobehavioural abnormalities by papaverine in prenatal valproic acid rat model of ASD. *European journal of pharmacology* **2021**, *890*, 173663, doi:10.1016/j.ejphar.2020.173663.
35. Jabarin, R.; Netser, S.; Wagner, S. Beyond the three-chamber test: toward a multimodal and objective assessment of social behavior in rodents. *Molecular autism* **2022**, *13*, 41, doi:10.1186/s13229-022-00521-6.
36. Lyu, T.J.; Ma, J.; Zhang, X.Y.; Xie, G.G.; Liu, C.; Du, J.; Xu, Y.N.; Yang, D.C.; Cen, C.; Wang, M.Y.; et al. Deficiency of FRMD5 results in neurodevelopmental dysfunction and autistic-like behavior in mice. *Mol Psychiatry* **2024**, doi:10.1038/s41380-024-02407-w.
37. Morel, C.; Paoli, J.; Camonin, C.; Marchal, N.; Grova, N.; Schroeder, H. Comparison of predictive validity of two autism spectrum disorder rat models: Behavioural investigations. *Neurotoxicology* **2024**, *103*, 39-49, doi:10.1016/j.neuro.2024.05.002.
38. Wilson, R.J.; Suh, Y.P.; Dursun, I.; Li, X.; da Costa Souza, F.; Grodzki, A.C.; Cui, J.Y.; Lehmler, H.J.; Lein, P.J. Developmental exposure to the Fox River PCB mixture modulates behavior in juvenile mice. *Neurotoxicology* **2024**, *103*, 146-161, doi:10.1016/j.neuro.2024.06.008.
39. Takumi, T.; Tamada, K.; Hatanaka, F.; Nakai, N.; Bolton, P.F. Behavioral neuroscience of autism. *Neuroscience and biobehavioral reviews* **2020**, *110*, 60-76, doi:10.1016/j.neubiorev.2019.04.012.
40. Silverman, J.L.; Yang, M.; Lord, C.; Crawley, J.N. Behavioural phenotyping assays for mouse models of autism. *Nature Reviews Neuroscience* **2010**, *11*, 490-502, doi:10.1038/nrn2851.
41. Kudryavtseva, N.N. Use of the "partition" test in behavioral and pharmacological experiments. *Neuroscience and behavioral physiology* **2003**, *33*, 461-471, doi:10.1023/a:1023411217051.
42. Lawande, N.V.; Ujjainwala, A.L.; Christian, C.A. A Single Test to Study Social Behavior and Repetitive Self-grooming in Mice. *Bio-protocol* **2020**, *10*, doi:10.21769/BioProtoc.3499.
43. Spencer, C.M.; Graham, D.F.; Yuva-Paylor, L.A.; Nelson, D.L.; Paylor, R. Social behavior in Fmr1 knockout mice carrying a human FMR1 transgene. *Behavioral neuroscience* **2008**, *122*, 710-715, doi:10.1037/0735-7044.122.3.710.
44. Elizabeth L. Berg; Jill L. Silverman. Animal models of autism. In *The Neuroscience of Autism*, Rajesh K. Kana, Ed.; Academic Press,: 2022; pp. 157-196,.
45. Hirsch, M.M.; Deckmann, I.; Fontes-Dutra, M.; Bauer-Negrini, G.; Nunes, G.D.-F.; Nunes, W.; Rabelo, B.; Riesgo, R.; Margis, R.; Bambini-Junior, V. Behavioral alterations in autism model induced by valproic acid and translational analysis of circulating microRNA. *Food and Chemical Toxicology* **2018**, *115*, 336-343.
46. d'Isa, R.; Fasano, S.; Brambilla, R. Editorial: Animal-friendly methods for rodent behavioral testing in neuroscience research. *Frontiers in behavioral neuroscience* **2024**, *18*, 1431310, doi:10.3389/fnbeh.2024.1431310.
47. Hrabovska, S.V.; Salyha, Y.T. Animal Models of Autism Spectrum Disorders and Behavioral Techniques of their Examination. *Neurophysiology* **2016**, *48*, 380-388, doi:10.1007/s11062-017-9613-2.
48. Brunner, D.; Kabitzke, P.; He, D.; Cox, K.; Thiede, L.; Hanania, T.; Sabath, E.; Alexandrov, V.; Saxe, M.; Peles, E.; et al. Comprehensive Analysis of the 16p11.2 Deletion and Null Cntnap2 Mouse Models of Autism Spectrum Disorder. *PLoS One* **2015**, *10*, e0134572, doi:10.1371/journal.pone.0134572.
49. Jones, K.L.; Pride, M.C.; Edmiston, E.; Yang, M.; Silverman, J.L.; Crawley, J.N.; Van de Water, J. Autism-specific maternal autoantibodies produce behavioral abnormalities in an endogenous antigen-driven mouse model of autism. *Mol Psychiatry* **2020**, *25*, 2994-3009, doi:10.1038/s41380-018-0126-1.
50. Moy, S.S.; Nadler, J.J.; Perez, A.; Barbaro, R.P.; Johns, J.M.; Magnuson, T.R.; Piven, J.; Crawley, J.N. Sociability and preference for social novelty in five inbred strains: an approach to assess autistic-like behavior in mice. *Genes, brain, and behavior* **2004**, *3*, 287-302, doi:10.1111/j.1601-1848.2004.00076.x.
51. Ranger, P. The development of animal models for autism: A gene-environment approach. **2016**, doi:10.26686/WGTN.17019029.V1.
52. Martin, L.; Sample, H.; Gregg, M.; Wood, C. Validation of operant social motivation paradigms using BTBR T+tf/J and C57BL/6J inbred mouse strains. *Brain and behavior* **2014**, *4*, 754-764, doi:10.1002/brb3.273.
53. Maloney, S.E.; Sarafinowska, S.; Weichselbaum, C.; McCullough, K.B.; Swift, R.G.; Liu, Y.; Dougherty, J.D. A comprehensive assay of social motivation reveals sex-specific roles of autism-associated genes and oxytocin. *Cell reports methods* **2023**, *3*, 100504, doi:10.1016/j.crmeth.2023.100504.

54. Premoli, M.; Pietropaolo, S.; Wöhr, M.; Simola, N.; Bonini, S.A. Mouse and rat ultrasonic vocalizations in neuroscience and neuropharmacology: State of the art and future applications. *European Journal of Neuroscience* **2023**, *n/a*, doi:https://doi.org/10.1111/ejn.15957.
55. Hahn, M.E.; Lavooy, M.J. A review of the methods of studies on infant ultrasound production and maternal retrieval in small rodents. *Behavior genetics* **2005**, *35*, 31-52, doi:10.1007/s10519-004-0854-7.
56. Stark, R.A.; Harker, A.; Salamanca, S.; Pellis, S.M.; Li, F.; Gibb, R.L. Development of ultrasonic calls in rat pups follows similar patterns regardless of isolation distress. *Dev Psychobiol* **2020**, *62*, 617-630, doi:10.1002/dev.21933.
57. Yamauchi, T.; Yoshioka, T.; Yamada, D.; Hamano, T.; Ohashi, M.; Matsumoto, M.; Iio, K.; Ikeda, M.; Kamei, M.; Otsuki, T.; et al. Cold-Restraint Stress-Induced Ultrasonic Vocalization as a Novel Tool to Measure Anxiety in Mice. *Biological and Pharmaceutical Bulletin* **2022**, *45*, 268-275, doi:10.1248/bpb.b21-00776.
58. Vadnie, C.A.; DePoy, L.M.; McClung, C.A. Measuring the Effects of Circadian Rhythm-Related Manipulations on Depression-Like Behavior in Rodents: Forced Swim and Tail Suspension Tests. *Methods in molecular biology (Clifton, N.J.)* **2021**, *2130*, 69-78, doi:10.1007/978-1-0716-0381-9_5.
59. Portfors, C.V. Types and functions of ultrasonic vocalizations in laboratory rats and mice. *Journal of the American Association for Laboratory Animal Science: JAALAS* **2007**, *46*, 28-34.
60. Van Segbroeck, M.; Knoll, A.T.; Levitt, P.; Narayanan, S. MUPET-Mouse Ultrasonic Profile ExTraction: A Signal Processing Tool for Rapid and Unsupervised Analysis of Ultrasonic Vocalizations. *Neuron* **2017**, *94*, 465-485.e465, doi:10.1016/j.neuron.2017.04.005.
61. Coffey, K.R.; Marx, R.E.; Neumaier, J.F. DeepSqueak: a deep learning-based system for detection and analysis of ultrasonic vocalizations. *Neuropsychopharmacology* **2019**, *44*, 859-868, doi:10.1038/s41386-018-0303-6.
62. Caruso, A.; Ricceri, L.; Scattoni, M.L. Ultrasonic vocalizations as a fundamental tool for early and adult behavioral phenotyping of Autism Spectrum Disorder rodent models. *Neuroscience and biobehavioral reviews* **2020**, *116*, 31-43, doi:10.1016/j.neubiorev.2020.06.011.
63. Gal, A.; Raykin, E.; Giladi, S.; Lederman, D.; Kofman, O.; Golan, H.M. Temporal dynamics of isolation calls emitted by pups in environmental and genetic mouse models of autism spectrum disorder. *Frontiers in neuroscience* **2023**, *17*, 1274039, doi:10.3389/fnins.2023.1274039.
64. Scattoni, M.L.; Ricceri, L.; Crawley, J.N. Unusual repertoire of vocalizations in adult BTBR T+tf/J mice during three types of social encounters. *Genes, brain, and behavior* **2011**, *10*, 44-56, doi:10.1111/j.1601-183X.2010.00623.x.
65. Sungur, A.; Schwarting, R.K.; Wöhr, M. Early communication deficits in the Shank1 knockout mouse model for autism spectrum disorder: Developmental aspects and effects of social context. *Autism research: official journal of the International Society for Autism Research* **2016**, *9*, 696-709, doi:10.1002/aur.1564.
66. Branchi, I.; Santucci, D.; Alleva, E. Ultrasonic vocalisation emitted by infant rodents: a tool for assessment of neurobehavioural development. *Behav Brain Res* **2001**, *125*, 49-56, doi:10.1016/s0166-4328(01)00277-7.
67. Scattoni, M.L.; Crawley, J.; Ricceri, L. Ultrasonic vocalizations: a tool for behavioural phenotyping of mouse models of neurodevelopmental disorders. *Neuroscience and biobehavioral reviews* **2009**, *33*, 508-515, doi:10.1016/j.neubiorev.2008.08.003.
68. Peça, J.; Feliciano, C.; Ting, J.T.; Wang, W.; Wells, M.F.; Venkatraman, T.N.; Lascola, C.D.; Fu, Z.; Feng, G. Shank3 mutant mice display autistic-like behaviours and striatal dysfunction. *Nature* **2011**, *472*, 437-442, doi:10.1038/nature09965.
69. Peñagarikano, O.; Abrahams, B.S.; Herman, E.I.; Winden, K.D.; Gdalyahu, A.; Dong, H.; Sonnenblick, L.I.; Gruver, R.; Almajano, J.; Bragin, A.; et al. Absence of CNTNAP2 leads to epilepsy, neuronal migration abnormalities, and core autism-related deficits. *Cell* **2011**, *147*, 235-246, doi:10.1016/j.cell.2011.08.040.
70. Möhrle, D.; Yuen, M.; Zheng, A.; Haddad, F.L.; Allman, B.L.; Schmid, S. Characterizing maternal isolation-induced ultrasonic vocalizations in a gene-environment interaction rat model for autism. *Genes, brain, and behavior* **2023**, *22*, e12841, doi:10.1111/gbb.12841.
71. Berg, E.L.; Copping, N.A.; Rivera, J.K.; Pride, M.C.; Careaga, M.; Bauman, M.D.; Berman, R.F.; Lein, P.J.; Harony-Nicolas, H.; Buxbaum, J.D.; et al. Developmental social communication deficits in the Shank3 rat

- model of phelan-mcdermid syndrome and autism spectrum disorder. *Autism research: official journal of the International Society for Autism Research* **2018**, *11*, 587-601, doi:10.1002/aur.1925.
72. Takahashi, T.; Okabe, S.; Broin, P.; Nishi, A.; Ye, K.; Beckert, M.V.; Izumi, T.; Machida, A.; Kang, G.; Abe, S.; et al. Structure and function of neonatal social communication in a genetic mouse model of autism. *Mol Psychiatry* **2016**, *21*, 1208-1214, doi:10.1038/mp.2015.190.
 73. Moskal, J.R.; Burgdorf, J. Chapter 39 - Ultrasonic Vocalizations in Rats as a Measure of Emotional Responses to Stress: Models of Anxiety and Depression. In *Handbook of Behavioral Neuroscience*, Brudzynski, S.M., Ed.; Elsevier: 2018; Volume 25, pp. 413-421.
 74. Chabout, J.; Serreau, P.; Ey, E.; Bellier, L.; Aubin, T.; Bourgeron, T.; Granon, S. Adult male mice emit context-specific ultrasonic vocalizations that are modulated by prior isolation or group rearing environment. *PLoS One* **2012**, *7*, e29401, doi:10.1371/journal.pone.0029401.
 75. Becker, M.; Gorobets, D.; Shmerkin, E.; Weinstein-Fudim, L.; Pinhasov, A.; Ornoy, A. Prenatal SAME Treatment Changes via Epigenetic Mechanism/s USVs in Young Mice and Hippocampal Monoamines Turnover at Adulthood in a Mouse Model of Social Hierarchy and Depression. *Int J Mol Sci* **2023**, *24*, doi:10.3390/ijms241310721.
 76. Boulay, D.; Ho-Van, S.; Bergis, O.; Avenet, P.; Griebel, G. Phencyclidine decreases tickling-induced 50-kHz ultrasound vocalizations in juvenile rats: a putative model of the negative symptoms of schizophrenia? *Behavioural pharmacology* **2013**, *24*, 543-551, doi:10.1097/FBP.0b013e3283654044.
 77. Nikiforuk, A.; Kos, T.; Fijał, K.; Hołuj, M.; Rafa, D.; Popik, P. Effects of the selective 5-HT7 receptor antagonist SB-269970 and amisulpride on ketamine-induced schizophrenia-like deficits in rats. *PLoS One* **2013**, *8*, e66695, doi:10.1371/journal.pone.0066695.
 78. Faure, A.; Nosjean, A.; Pittaras, E.; Duchêne, A.; Andrieux, A.; Gory-Fauré, S.; Charvériat, M.; Granon, S. Dissociated features of social cognition altered in mouse models of schizophrenia: Focus on social dominance and acoustic communication. *Neuropharmacology* **2019**, *159*, 107334, doi:10.1016/j.neuropharm.2018.09.009.
 79. Gandal, M.J.; Anderson, R.L.; Billingslea, E.N.; Carlson, G.C.; Roberts, T.P.; Siegel, S.J. Mice with reduced NMDA receptor expression: more consistent with autism than schizophrenia? *Genes, brain, and behavior* **2012**, *11*, 740-750, doi:10.1111/j.1601-183X.2012.00816.x.
 80. Sterling, M.L.; Teunisse, R.; Englitz, B. Rodent ultrasonic vocal interaction resolved with millimeter precision using hybrid beamforming. *eLife* **2023**, *12*, doi:10.7554/eLife.86126.
 81. Boulanger-Bertolus, J.; Mouly, A.M. Ultrasonic Vocalizations Emission across Development in Rats: Coordination with Respiration and Impact on Brain Neural Dynamics. *Brain Sci* **2021**, *11*, doi:10.3390/brainsci11050616.
 82. Ou-Yang, B.; Hu, Y.; Fei, X.Y.; Cheng, S.T.; Hang, Y.; Yang, C.; Cheng, L. A meta-analytic study of the effects of early maternal separation on cognitive flexibility in rodent offspring. *Developmental cognitive neuroscience* **2022**, *56*, 101126, doi:10.1016/j.dcn.2022.101126.
 83. Cowan, C.S.M.; Richardson, R. A Brief Guide to Studying Fear in Developing Rodents: Important Considerations and Common Pitfalls. *Current protocols in neuroscience* **2018**, *83*, e44, doi:10.1002/cpns.44.
 84. Kabitzke, P.; Morales, D.; He, D.; Cox, K.; Sutphen, J.; Thiede, L.; Sabath, E.; Hanania, T.; Biemans, B.; Brunner, D. Mouse model systems of autism spectrum disorder: Replicability and informatics signature. *Genes, brain, and behavior* **2020**, *19*, e12676, doi:10.1111/gbb.12676.
 85. Wöhr, M.; Rouillet, F.I.; Crawley, J.N. Reduced scent marking and ultrasonic vocalizations in the BTBR T+tf/J mouse model of autism. *Genes, brain, and behavior* **2011**, *10*, 35-43, doi:10.1111/j.1601-183X.2010.00582.x.
 86. Wöhr, M.; Rouillet, F.I.; Hung, A.Y.; Sheng, M.; Crawley, J.N. Communication impairments in mice lacking Shank1: reduced levels of ultrasonic vocalizations and scent marking behavior. *PLoS One* **2011**, *6*, e20631, doi:10.1371/journal.pone.0020631.
 87. Basurto, E.; González-Flores, O.; Hoffman, K. Chronic MK-801 administration provokes persistent deficits in social memory in the prairie vole (*Microtus ochrogaster*): A potential animal model for social deficits of schizophrenia. *Behav Brain Res* **2024**, *465*, 114948, doi:10.1016/j.bbr.2024.114948.

88. Malik, J.A.; Yaseen, Z.; Thotapalli, L.; Ahmed, S.; Shaikh, M.F.; Anwar, S. Understanding translational research in schizophrenia: A novel insight into animal models. *Molecular biology reports* **2023**, *50*, 3767-3785, doi:10.1007/s11033-023-08241-7.
89. Yao, Y.; Jin, C.; Liao, Y.; Huang, X.; Wei, Z.; Zhang, Y.; Li, D.; Su, H.; Han, W.; Qin, D. Schizophrenia-Like Behaviors Arising from Dysregulated Proline Metabolism Are Associated with Altered Neuronal Morphology and Function in Mice with Hippocampal PRODH Deficiency. *Aging and disease* **2024**, *15*, 1952-1968, doi:10.14336/ad.2023.0902.
90. Eshraghi, A.A.; Memis, I.; Wang, F.; White, I.; Furar, E.; Mittal, J.; Moosa, M.; Atkins, C.M.; Mittal, R. Genetic ablation of metabotropic glutamate receptor 5 in rats results in an autism-like behavioral phenotype. *PLoS One* **2022**, *17*, e0275937, doi:10.1371/journal.pone.0275937.
91. de Brouwer, G.; Fick, A.; Harvey, B.H.; Wolmarans, W. A critical inquiry into marble-burying as a preclinical screening paradigm of relevance for anxiety and obsessive-compulsive disorder: Mapping the way forward. *Cognitive, affective & behavioral neuroscience* **2019**, *19*, 1-39, doi:10.3758/s13415-018-00653-4.
92. Dunn, J.T.; Guidotti, A.; Grayson, D.R. Behavioral and Molecular Characterization of Prenatal Stress Effects on the C57BL/6J Genetic Background for the Study of Autism Spectrum Disorder. *eNeuro* **2024**, *11*, doi:10.1523/eneuro.0186-23.2024.
93. Jones, C.A.; Watson, D.J.; Fone, K.C. Animal models of schizophrenia. *British journal of pharmacology* **2011**, *164*, 1162-1194, doi:10.1111/j.1476-5381.2011.01386.x.
94. Carmel, J.; Ghanayem, N.; Mayouf, R.; Saleev, N.; Chatterjee, I.; Getselter, D.; Tikhonov, E.; Turjeman, S.; Shaalan, M.; Khateeb, S.; et al. Bacteroides is increased in an autism cohort and induces autism-relevant behavioral changes in mice in a sex-dependent manner. *NPJ biofilms and microbiomes* **2023**, *9*, 103, doi:10.1038/s41522-023-00469-2.
95. Sato, A.; Ikeda, K. Genetic and Environmental Contributions to Autism Spectrum Disorder Through Mechanistic Target of Rapamycin. *Biological psychiatry global open science* **2022**, *2*, 95-105, doi:10.1016/j.bpsgos.2021.08.005.
96. Ornoy, A.; Weinstein-Fudim, L.; Tfilin, M.; Ergaz, Z.; Yanai, J.; Szyf, M.; Turgeman, G. S-adenosyl methionine prevents ASD like behaviors triggered by early postnatal valproic acid exposure in very young mice. *Neurotoxicology and Teratology* **2019**, *71*, 64-74, doi:https://doi.org/10.1016/j.ntt.2018.01.005.
97. Rodríguez Peris, L.; Scheuber, M.I.; Shan, H.; Braun, M.; Schwab, M.E. Barnes maze test for spatial memory: A new, sensitive scoring system for mouse search strategies. *Behav Brain Res* **2024**, *458*, 114730, doi:10.1016/j.bbr.2023.114730.
98. Dougnon, G.; Matsui, H. Modelling Autism Spectrum Disorder (ASD) and Attention-Deficit/Hyperactivity Disorder (ADHD) Using Mice and Zebrafish. *Int J Mol Sci* **2022**, *23*, doi:10.3390/ijms23147550.
99. Li, L.; Nguyen, A.; Luo, J. Mouse model of repetitive mild traumatic brain injury (investigating cognitive impairment). In *Handbook of Animal Models in Neurological Disorders*; Elsevier: 2023; pp. 121-132.
100. Ait Bali, Y.; Madougou, F.A.; Ba-M'hamed, S.; Giustetto, M.; Bennis, M. Glyphosate-based herbicide exposure affects cognitive flexibility and social cognition in adult mice. *Neuroscience letters* **2024**, *837*, 137912, doi:10.1016/j.neulet.2024.137912.
101. Avraham, Y.; Berry, E.M.; Donskoy, M.; Ahmad, W.A.; Vorobiev, L.; Albeck, A.; Mankuta, D. Beta-carotene as a novel therapy for the treatment of "Autistic like behavior" in animal models of Autism. *Behav Brain Res* **2019**, *364*, 469-479, doi:10.1016/j.bbr.2017.09.041.
102. Hirsch, M.M.; Deckmann, I.; Santos-Terra, J.; Staevie, G.Z.; Fontes-Dutra, M.; Carello-Collar, G.; Körbes-Rockenbach, M.; Brum Schwingel, G.; Bauer-Negrini, G.; Rabelo, B.; et al. Effects of single-dose antipurinergic therapy on behavioral and molecular alterations in the valproic acid-induced animal model of autism. *Neuropharmacology* **2020**, *167*, 107930, doi:10.1016/j.neuropharm.2019.107930.
103. Brownstien, M.; Lazar, M.; Botvinnik, A.; Shevakh, C.; Blakolmer, K.; Lerer, L.; Lifschytz, T.; Lerer, B. Striking Long Term Beneficial Effects of Single Dose Psilocybin and Psychedelic Mushroom Extract in the SAPAP3 Rodent Model of OCD-Like Excessive Self-Grooming. *bioRxiv* **2024**, doi:10.1101/2024.06.25.600634.
104. Wu, Y.; Qi, F.; Song, D.; He, Z.; Zuo, Z.; Yang, Y.; Liu, Q.; Hu, S.; Wang, X.; Zheng, X.; et al. Prenatal influenza vaccination rescues impairments of social behavior and lamination in a mouse model of autism. *Journal of neuroinflammation* **2018**, *15*, 228, doi:10.1186/s12974-018-1252-z.

105. Xiao, L.; Jiang, S.; Wang, Y.; Gao, C.; Liu, C.; Huo, X.; Li, W.; Guo, B.; Wang, C.; Sun, Y.; et al. Continuous high-frequency deep brain stimulation of the anterior insula modulates autism-like behavior in a valproic acid-induced rat model. *Journal of translational medicine* **2022**, *20*, 570, doi:10.1186/s12967-022-03787-9.
106. Jamwal, S.; Islam, R.; Kaswan, Z.M.; Ahmed, S.; Bowers, C.; Giuliano, L.; Kaffman, A. Postnatal Enrichment Corrects Deficits in Perineuronal Net Formation and Reversal Learning in Adult Mice Exposed to Early Adversity. *bioRxiv* **2024**, doi:10.1101/2024.09.21.614236.
107. Zhang, Y.; Xiang, Z.; Jia, Y.; He, X.; Wang, L.; Cui, W. The Notch signaling pathway inhibitor Dapt alleviates autism-like behavior, autophagy and dendritic spine density abnormalities in a valproic acid-induced animal model of autism. *Progress in neuro-psychopharmacology & biological psychiatry* **2019**, *94*, 109644, doi:10.1016/j.pnpbp.2019.109644.
108. Sharma, S.; Cheema, M.; Reeson, P.L.; Narayana, K.; Boghazian, R.; Cota, A.P.; Brosschot, T.P.; FitzPatrick, R.D.; Körbelin, J.; Reynolds, L.A.; et al. A pathogenic role for IL-10 signalling in capillary stalling and cognitive impairment in type 1 diabetes. *Nature metabolism* **2024**, *6*, 2082-2099, doi:10.1038/s42255-024-01159-9.
109. AlQot, H.E.; Rylett, R.J. Primate-specific 82-kDa choline acetyltransferase attenuates progression of Alzheimer's disease-like pathology in the APP(NL-G-F) knock-in mouse model. *Scientific reports* **2024**, *14*, 27614, doi:10.1038/s41598-024-78751-2.
110. Elam, H.B.; Perez, S.M.; Donegan, J.J.; Eassa, N.E.; Lodge, D.J. Knockdown of Lhx6 during embryonic development results in neurophysiological alterations and behavioral deficits analogous to schizophrenia in adult rats. *Schizophr Res* **2024**, *267*, 113-121, doi:10.1016/j.schres.2024.03.032.
111. Young, J.W.; Zhou, X.; Geyer, M.A. Animal models of schizophrenia. *Current topics in behavioral neurosciences* **2010**, *4*, 391-433, doi:10.1007/7854_2010_62.
112. Korkmaz, F.; Sims, S.; Sen, F.; Sultana, F.; Laurencin, V.; Cullen, L.; Pallapati, A.; Liu, A.; Chen, R.; Rojekar, S.; et al. Gene-dose-dependent reduction of Fshr expression improves spatial memory deficits in Alzheimer's mice. *Mol Psychiatry* **2024**, doi:10.1038/s41380-024-02824-x.
113. Miao, Z.; Chen, L.; Zhang, Y.; Zhang, J.; Zhang, H. Bifidobacterium animalis subsp. lactis Probio-M8 alleviates abnormal behavior and regulates gut microbiota in a mouse model suffering from autism. *mSystems* **2024**, *9*, e0101323, doi:10.1128/msystems.01013-23.
114. Taskiran, M.; Yildiz Taskiran, S.; Unal, G.; Bozkurt, N.M.; Golgeli, A. Vortioxetine improved schizophrenia-like behavioral deficits in a Poly I:C-induced maternal immune activation model of schizophrenia in rats. *Fundamental & clinical pharmacology* **2024**, *38*, 1069-1079, doi:10.1111/fcp.13028.
115. Markram, K.; Rinaldi, T.; La Mendola, D.; Sandi, C.; Markram, H. Abnormal fear conditioning and amygdala processing in an animal model of autism. *Neuropsychopharmacology* **2008**, *33*, 901-912, doi:10.1038/sj.npp.1301453.
116. Lanjewar, A.L.; Levitt, P.; Eagleson, K.L. Developmental and molecular contributions to contextual fear memory emergence in mice. *Neuropsychopharmacology* **2024**, *49*, 1392-1401, doi:10.1038/s41386-024-01835-0.
117. Zhang, W.; Huang, J.; Gao, F.; You, Q.; Ding, L.; Gong, J.; Zhang, M.; Ma, R.; Zheng, S.; Sun, X.; et al. Lactobacillus reuteri normalizes altered fear memory in male Cntnap4 knockout mice. *EBioMedicine* **2022**, *86*, 104323, doi:10.1016/j.ebiom.2022.104323.
118. Lalonde, R. The neurobiological basis of spontaneous alternation. *Neuroscience and biobehavioral reviews* **2002**, *26*, 91-104, doi:10.1016/s0149-7634(01)00041-0.
119. Li, Z.; Zhu, Y.X.; Gu, L.J.; Cheng, Y. Understanding autism spectrum disorders with animal models: applications, insights, and perspectives. *Zoological research* **2021**, *42*, 800-824, doi:10.24272/j.issn.2095-8137.2021.251.
120. Fereshetyan, K.; Chavushyan, V.; Danielyan, M.; Yenkovyan, K. Assessment of behavioral, morphological and electrophysiological changes in prenatal and postnatal valproate induced rat models of autism spectrum disorder. *Scientific Reports* **2021**, *11*, 23471, doi:10.1038/s41598-021-02994-6.
121. Wang, H.; Wang, Q.; Cui, L.; Feng, X.; Dong, P.; Tan, L.; Lin, L.; Lian, H.; Cao, S.; Huang, H.; et al. A molecularly defined amygdala-independent tetra-synaptic forebrain-to-hindbrain pathway for odor-driven innate fear and anxiety. *Nature neuroscience* **2024**, *27*, 514-526, doi:10.1038/s41593-023-01562-7.

122. Sharma, A.R.; Batra, G.; Dhir, N.; Jain, A.; Modi, T.; Saini, L.; Thakur, N.; Mishra, A.; Singh, R.S.; Singh, A.; et al. "Comparative evaluation of different chemical agents induced Autism Spectrum Disorder in experimental Wistar rats". *Behav Brain Res* **2024**, *458*, 114728, doi:10.1016/j.bbr.2023.114728.
123. Yan, L.; Yang, F.; Wang, Y.; Shi, L.; Wang, M.; Yang, D.; Wang, W.; Jia, Y.; So, K.F.; Zhang, L. Stress increases hepatic release of lipocalin 2 which contributes to anxiety-like behavior in mice. *Nature communications* **2024**, *15*, 3034, doi:10.1038/s41467-024-47266-9.
124. Zhang, G.W.; Shen, L.; Tao, C.; Jung, A.H.; Peng, B.; Li, Z.; Zhang, L.I.; Tao, H.W. Medial preoptic area antagonistically mediates stress-induced anxiety and parental behavior. *Nature neuroscience* **2021**, *24*, 516-528, doi:10.1038/s41593-020-00784-3.
125. Neuwirth, L.S.; Verrengia, M.T.; Harikinish-Murray, Z.I.; Orens, J.E.; Lopez, O.E. Under or Absent Reporting of Light Stimuli in Testing of Anxiety-Like Behaviors in Rodents: The Need for Standardization. *Frontiers in molecular neuroscience* **2022**, *15*, 912146, doi:10.3389/fnmol.2022.912146.
126. Lubrich, C.; Giesler, P.; Kipp, M. Motor Behavioral Deficits in the Cuprizone Model: Validity of the Rotarod Test Paradigm. *Int J Mol Sci* **2022**, *23*, doi:10.3390/ijms231911342.
127. Keane, S.P.; Chadman, K.K.; Gomez, A.R.; Hu, W. Pros and cons of narrow- versus wide-compartment rotarod apparatus: An experimental study in mice. *Behav Brain Res* **2024**, *463*, 114901, doi:10.1016/j.bbr.2024.114901.
128. Deacon, R.M. Measuring motor coordination in mice. *Journal of visualized experiments : JoVE* **2013**, e2609, doi:10.3791/2609.
129. Bidgood, R.; Zubelzu, M.; Ruiz-Ortega, J.A.; Morera-Herreras, T. Automated procedure to detect subtle motor alterations in the balance beam test in a mouse model of early Parkinson's disease. *Scientific reports* **2024**, *14*, 862, doi:10.1038/s41598-024-51225-1.
130. Ito, S.; Kakuta, Y.; Yoshida, K.; Shiota, Y.; Mieda, T.; Iizuka, Y.; Chikuda, H.; Iizuka, H.; Nakamura, K. A simple scoring of beam walking performance after spinal cord injury in mice. *PLoS One* **2022**, *17*, e0272233, doi:10.1371/journal.pone.0272233.
131. Cording, K.R.; Bateup, H.S. Altered motor learning and coordination in mouse models of autism spectrum disorder. *Frontiers in cellular neuroscience* **2023**, *17*, 1270489, doi:10.3389/fncel.2023.1270489.
132. Guilloux, J.-P.; Seney, M.; Edgar, N.; Sibille, E. Integrated behavioral z-scoring increases the sensitivity and reliability of behavioral phenotyping in mice: Relevance to emotionality and sex. *Journal of Neuroscience Methods* **2011**, *197*, 21-31, doi:https://doi.org/10.1016/j.jneumeth.2011.01.019.
133. Kraeuter, A.-K. The use of integrated behavioural z-scoring in behavioural neuroscience – A perspective article. *Journal of Neuroscience Methods* **2023**, *384*, 109751, doi:https://doi.org/10.1016/j.jneumeth.2022.109751.
134. Labots, M.; Laarakker, M.C.; Schetters, D.; Arndt, S.S.; van Lith, H.A. An improved procedure for integrated behavioral z-scoring illustrated with modified Hole Board behavior of male inbred laboratory mice. *Journal of Neuroscience Methods* **2018**, *293*, 375-388, doi:https://doi.org/10.1016/j.jneumeth.2017.09.003.
135. El-Kordi, A.; Winkler, D.; Hammerschmidt, K.; Kästner, A.; Krueger, D.; Ronnenberg, A.; Ritter, C.; Jatho, J.; Radyushkin, K.; Bourgeron, T.; et al. Development of an autism severity score for mice using *Nlgn4* null mutants as a construct-valid model of heritable monogenic autism. *Behav Brain Res* **2013**, *251*, 41-49, doi:10.1016/j.bbr.2012.11.016.
136. Ornoy, A.; Echefu, B.; Becker, M. Valproic Acid in Pregnancy Revisited: Neurobehavioral, Biochemical and Molecular Changes Affecting the Embryo and Fetus in Humans and in Animals: A Narrative Review. *Int J Mol Sci* **2023**, *25*, doi:10.3390/ijms25010390.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.