Department for Biomedical Science
University of Veterinary Medicine Vienna

Institute of Laboratory Animal Science
(Head: Univ.-Prof. Dr. Maik Dahlhoff)

Review of Experimental Conditions in Phenotyping Studies in Laboratory Mice with Regard to Severity Assessment According to the TGV2012

Diploma Thesis

For Obtaining the Academic Degree

Magistrae Medicinae Veterinariae

University of Veterinary Medicine

submitted by
Jasmin Keller

Vienna, November 2020

Academic supervisor: Ao.Univ.-Prof. Dipl.-Ing. Dr.sc.agr. Thomas Kolbe

First assessor: Ao.Univ.-Prof. Dipl.-Ing. Dr.sc.agr. Thomas Kolbe

Second assessor: Univ.-Prof. Jean-Loup Rault, PhD

Table of Contents

1. Introduction 1
2. Search Protocol 4
   2.1. Subject Analysis 4
   2.2. Search Keys 4
   2.3. Search Tools 5
   2.4. Accessed Data 5
3. Behavioural Tests 7
4. General Health 8
   4.1. Abnormal Spontaneous Behaviour 8
   4.2. Activity Rhythms, Sleep and Home Cage Activity Patterns 8
   4.3. Appearance in General 9
   4.4. Body and Skin Temperature 9
   4.5. Body Weight 10
   4.6. Neurological Reflexes 10
   4.7. Posture and Gait 10
   4.8. Seizures and Vestibular Dysfunction 11
   4.9. Sensory Capabilities 11
5. Classification of Behavioural Tests 12
   5.1. Motor Function 12
      5.1.1. Gait Analysis and Locomotor Function 12
         5.1.1.1. Circling 12
         5.1.1.2. Cylinder Test 12
         5.1.1.3. Foot Fault Test 13
         5.1.1.4. Footprint Pattern/CatWalk™ XT 14
         5.1.1.5. Rotometer 15
         5.1.1.6. Treadmill/DigiGait™ Test 16
      5.1.2. Locomotor Activity and Exploratory Behaviour 16
         5.1.2.1. Circadian Wheel Running 16
         5.1.2.2. Hole Poke Exploratory Test 16
         5.1.2.3. Open Field Test 18
      5.1.3. Motor Coordination and Balance 19
         5.1.3.1. Basket Test 19
         5.1.3.2. Beam Walking Test/Balance Beam Test 19
         5.1.3.3. Rotarod Test 21
         5.1.3.4. Staircase Test 23

5.1.3.5. Vertical Pole Test 23
5.1.4. Muscular Strength 23
5.1.4.1. Grip Strength Test 23
5.1.4.2. Wire Hang Test 25
5.2. Sensory Abilities 26
5.2.1. Olfactory Acuity and Olfactory Discrimination 27
5.2.2. Visual Acuity 28
5.2.2.1. Light-Dark Test 28
5.2.2.2. Visual Cliff Assay 30
5.2.2.3. Visual Water Box 31
5.2.3. Auditory Acuity and Acoustic Startle Response 31
5.2.4. Taste Acuity 31
5.2.5. Tactile Acuity 32
5.2.5.1. Air Puff Flinch Response 32
5.2.5.2. T/Y Maze 32
5.2.5.3. Von Frey Test 32
5.2.6. Nociceptive Sensitivity 32
5.2.6.1. Flinch Test 33
5.2.6.2. Formalin Test 33
5.2.6.3. Hot Plate Test 34
5.2.6.4. Plantar (Hargreaves) Test 36
5.2.6.5. Randall-Selitto Test/Paw-Pressure Test 36
5.2.6.6. Tail Flick Test 37
5.2.6.7. Tail Immersion/Hot-Water Tail Flick/Cold Water Tail Flick Test 38
5.2.6.8. Writhing Test 39
5.3. Learning and Memory 39
5.3.1. Spatial Learning 39
5.3.1.1. Barnes Maze 39
5.3.1.2. Cheese Board Test 42
5.3.1.3. Delayed Matching-To-Position and Delayed Non-Matching To-Position Performance 42
5.3.1.4. Morris Water Maze 42
5.3.1.5. T-Maze/Y-Maze Delayed Alternation 45
5.3.2. Associative Learning 45
5.3.2.1. Active Avoidance 45
5.3.2.2. Conditioned Fear 45
5.3.2.3. Conditioned Taste Aversion 49

5.3.2.4. Olfactory Discrimination 49
5.3.2.5. Passive Avoidance 49
5.3.2.6. Radial Maze 50
5.3.2.7. Step-Down Avoidance 50
5.3.2.8. Y-Maze/ T-Maze Avoidance 50
5.3.3. Recognition Memory 50
5.3.3.1. Social Interaction/Recognition Test 50
5.3.3.2. Object Recognition Test 52
5.3.4. Operant Learning 52
5.3.4.1. Five-Choice Serial Reaction Time Attentional Task 52
5.3.4.2. Schedule-Induced Operant Tasks 53
5.3.5. Motor Learning 53
5.3.5.1. Conditioning of Eyeblink 53
5.3.5.2. Rotarod 53
5.4. Feeding and Drinking 53
5.4.1. Analysis of Feeding Behaviour 54
5.4.2. Cafeteria Diet 54
5.4.3. Conditioned Taste Aversion 56
5.4.4. Conditioned Taste Preference 56
5.4.5. Restricted Daily Access 56
5.4.6. Specialised Diets 56
5.4.7. Twenty-Four Hour Consumption 57
5.5. Reproductive Behaviours 57
5.5.1. Courtship and Copulation 57
5.5.2. Nesting (Nest Building) 57
5.5.3. Parenting 58
5.5.4. Pup Retrieval 59
5.5.5. Pup Ultrasonic Vocalisation 59
5.6. Social Behaviours 59
5.6.1. Nesting (Nest building) 59
5.6.2. Social Interaction 59
5.6.2.1. Social Transmission of Food Preference 59
5.6.2.2. Three Chamber Social Approach Task 60
5.6.3. Aggression 60
5.6.3.1. Standard Opponent 60
5.6.3.2. Isolation Induced Fighting 60
5.6.3.3. Resident-Intruder Test 61

5.6.3.4. Round-Robin Approach 61
5.6.3.5. Social Dominance 61
5.7. Emotional Behaviours 61
5.7.1. Anxiety- and Fear-Related Behaviours 61
5.7.1.1. Conditioned Fear 61
5.7.1.2. Elevated Platform 62
5.7.1.3. Elevated Plus-Maze 62
5.7.1.4. Elevated Zero-Maze 62
5.7.1.5. Emergence Test 62
5.7.1.6. Fear/Defence Test Battery 63
5.7.1.7. Fear-Potentiated Startle 63
5.7.1.8. Four Plate Test 63
5.7.1.9. Learned Safety/Fear Inhibition 63
5.7.1.10. Light-Dark Test 64
5.7.1.11. Novelty Supressed Feeding 64
5.7.1.12. Open Field 64
5.7.1.13. Predator Odour 64
5.7.1.14. Pup Ultrasonic Vocalisations 64
5.7.1.15. Social Interaction 65
5.7.1.16. Visible Burrow System 65
5.7.1.17. Vogel Conflict Test 65
5.7.2. Behaviours and Models Related to Symptoms of Psychological Diseases 65
5.7.2.1. Marble Burying 65
5.7.2.2. Nestlet Shredding 66
5.7.3. Models of Schizophrenia 66
5.7.3.1. Acoustic Prepulse Inhibition 66
5.7.3.2. Behavioural Sensitisation 66
5.7.3.3. Latent Inhibition 66
5.7.3.4. Hyperlocomotion 67
5.7.4. Models of Depression 67
5.7.4.1. Immobilisation/Restraint 67
5.7.4.2. Forced Swim Test 68
5.7.4.3. Learned Helplessness 69
5.7.4.4. Postweaning Social Isolation 69
5.7.4.5. Social Deprivation/Social Isolation 69
5.7.4.6. Sucrose Preference Test 69

5.7.4.7. Tail Suspension Test 70
5.7.4.8. White Noise 70
5.8. Reward and Abuse 70
5.8.1. Conditioned Place Preference/Aversion 70
5.8.2. Intra Venous Self-Administration 70
5.8.3. Oral Self-Administration 71
5.8.4. Scoring Withdrawal Symptoms 71
6. Suggested Severity Assessment 72
7. Summary and Conclusion 75
8. Zusammenfassung 76
9. References 78
10. List of Figures and Tables 112
11. Abbreviations 113

1. Introduction

Within all animals used in research, mice are the most commonly used laboratory animals in Austria (Fig. 1) (Bundesministerium für Bildung, Wissenschaft und Forschung) and worldwide. On one hand, outbred strains are used in experiments, which represents human genetic diversity. Moreover, and more frequently, inbred strains are used due to their genetic stability, which improves the reproducibility of experiments (Lindzey and Delbert 1970). On the other hand, genetically modified mice have become inestimable to gain fundamental knowledge in basic and applied research, relate functions to genes and investigate numerous genetic disorders (Bockamp et al. 2002). Additionally, these mice are widely considered as model organisms for the investigation of human behavioural and neurological disorders, ranging from schizophrenia over depression to alcohol intoxication (Fox 2007).

As the outcome of a new genetic modification is often unknown, a series of different tests are required to phenotype the created strains. The results are compared with wildtype mice of the same strain, which must go through the same test protocol. This also includes behavioural testing to assess abnormal behaviour. There are more than 100 behavioural tests worldwide, which might differ in protocol and equipment, as there is no standard protocol specified (Crawley 2000), although there is an effort by 19 research institutions which are linked in the International Mouse Phenotyping Consortium (IMPC) to use identical test settings.

To use animals for research raised ethical concerns since centuries, beginning in the 18th century with Jeremy Bentham and since then with steadily increasing pressure by scientific and lay communities (Jeremy Bentham 1823, Sherwin et al. 2003). The Directive 2010/63/EU of the European Parliament on the protection of animals for scientific purposes intervenes expresses this concern and gives the legal framework for the scientific work with live vertebrate animals, cephalopods and larvae (European parliament and council of 22 September 2010 2010). It always strives for the principles of replacement, reduction and refinement. Beside these regulations it also stipulates all procedures to be assigned to a severity category, ranging from “non recovery” through “mild“ and “moderate“ to “severe“. This classification should not only help to balance the harm and benefits of an experiment, but also to reduce animal suffering and pain. Nevertheless, this classification is still based on theoretic considerations and subjective observations. There is limited informative value to assess and classify procedures given in the examples of severity classification in Annex VIII of the directive since they are little descriptive and limited in number.

The aim of this diploma thesis is to review publications describing phenotyping procedures in mice with emphasis on details of the experimental setup which could imply a variation of the burden of the tested animals.

According to the information given in these papers a range or a certain grade of severity is suggested in this thesis for every category of phenotyping experiments. This grading is not based on the biological stress occurring during instinctive behaviours like pup retrieval but on the juridical grading of the Austrian Tierversuchsgesetz 2012 (TVG2012). There are only four categories listed in the law: “terminal”, “mild”, “moderate”, “severe” and – of course, if nothing of this is applicable – “no burden”. The threshold from “no burden” to “mild” is defined as a professionally applied injection. In phenotyping test, there is commonly no permanent harm attributed to the mice. Therefore, the category “terminal” has to be excluded. Cancer models like Trp53 mice and comparable strains, immunological studies with Listeria monocytogenes infection for survival experiments, or partial or complete removal of vital organs is certainly classified as “severe” burden. Compared to this kind of studies phenotyping experiments ranges mostly far below this burden in the categories “mild” or “moderate”, if there is a burden according to the definition of the TVG2012 at all.

In November 2019 the Federal Ministry of Education, Science and Research published the recommendation “Stellungnahme zur Schweregradbeurteilung von Verhaltensexperimenten mit Mäusen und Ratten” (https://www.bmbwf.gv.at/Themen/Forschung/Forschung-in-%C3%96sterreich/Services/Tierversuche.html). This official recommendation was used as guideline to evaluate the influence of the experimental settings on animal welfare and suggest gradings for the different phenotyping tests reviewed in this thesis.

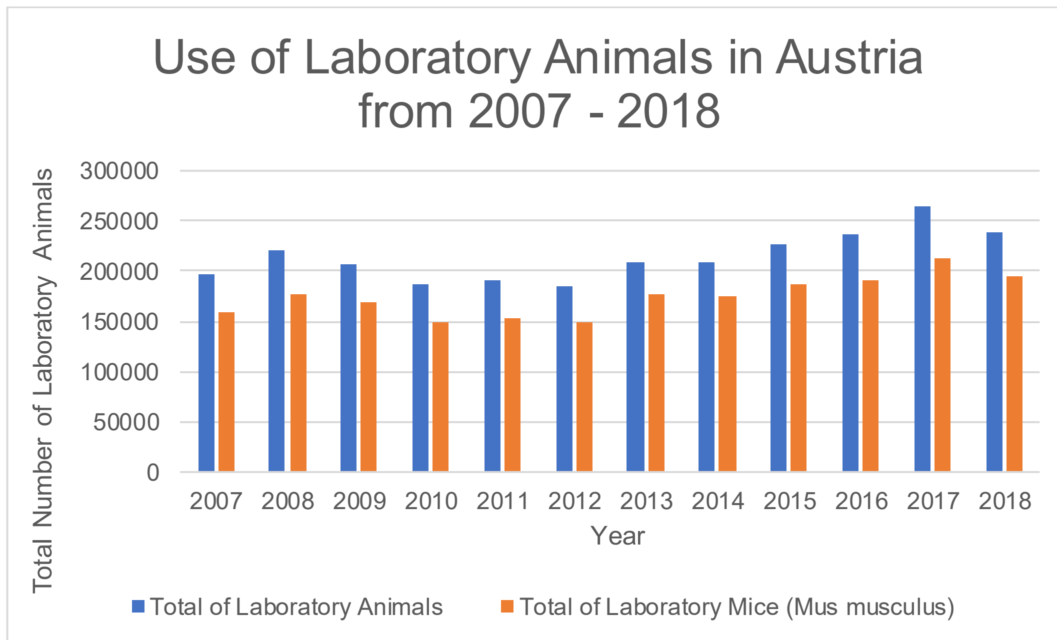

Figure 1: Use of laboratory animals in Austria from 2007 – 2018
Data from https://www.bmbwf.gv.at/Themen/Forschung/Forschung-in-Österreich/Services/Tierversuche/Tierversuchsstatistik.html (accessed 15.01.2020)

2. Search Protocol

The following protocol should give a short overview of the used research methods and help with documenting information gaining and structuring ideas. Therefore, we did a subject analysis, identified search keys and research tools and specified the accessed data.

2.1. Subject Analysis

Subject:

Laboratory mice are the most commonly used laboratory animals and behavioural tests are a deep-seated procedure in basic research. However, the classification of research procedures in the European Directive 2010/63/EU is still based on theoretic considerations and subjective observations.

Aim:

Review of publications describing phenotyping procedures in mice with emphasis on details of the experimental setup which could imply a variation of the burden of the tested animals. Additionally, a range of grades or a certain grade of severity was suggested for each category of phenotyping tests.

Terms, definitions and facts that have to be clarified or explained within this thesis as well as additional important facts could be the European Directive 2010/63/EU; factors which could be used for severity assessment; factors causing stress, fear and anxiety in mice; factors causing stress relaxation and reducing fear and anxiety in mice; animal welfare; most behavioural tests which are used for scientific purposes; the differences between used protocols of the behavioural tests.

2.2. Search Keys

Table 1: Keywords used for literature research

	Keywords			
Core Terms/Major Terms	Behavioural test	Stress	Stress relaxation	Laboratory mice, laboratory animals
Synonyms	Experiment	Anxiety, fear, nervousness	Comfort	Mice, Mouse

Generic Terms				Mus musculus
Subordinated Terms	All different tests			All different strains of laboratory mice
Related Terms	Test procedures, animal testing	panic, to be afraid, to be frightened, depression	Habituation, animal welfare	Laboratory rats, mouse strains

2.3. Search Tools

Institutions, websites or searching machines that could give relevant information about the topic and their accessibility are listed in the following table.

Table 2: Tools used for literature research

Institution/Website/Searching Machine	Access
Google Scholar	https://scholar.google.com
PubMed	https://pubmed.ncbi.nlm.nih.gov
Science Direct	https://www.sciencedirect.com
Nature	https://www.nature.com
Universitätsbibliothek Vetmeduni Vienna	https://www.vetmeduni.ac.at/bibliothek/
Elektronische Zeitschriftenbibliothek Universitätsbibliothek der Vetmeduni Vienna	http://ezb.uni-regensburg.de/?bibid=VMUW
Scopus	https://www.scopus.com/
Google	https://www.google.com

2.4. Accessed Data

First, we searched to identify as many used behavioural tests as possible. Therefore, we used the key words behavioural tests in mice; behavioural tests; behavioural testing mice; mice in research; mice behaviour and laboratory mice. After taking superficial notes of the different behavioural tests, we searched the name of the test to get particular and more detailed information about the individual tests. Additionally, a literature search using different research tools (Tab. 2) was performed to identify equal or different protocols of behavioural

tests used in research. We selected protocols for most recent studies, relevance, number of citations and mice used as testing animals. At last, we searched for the terms stress; animal welfare; fear; anxiety; stress in mice; stress laboratory mice or stress in connection to the particular behavioural test name, for underlining arguments in the discussion regarding the severity of the particular tests.

3. Behavioural Tests

Behavioural tests are used to study the behaviour of experimental animals in certain situations and under various conditions. This affects, for example, the investigation of the neurobiological basis of behaviour, the understanding of different disorders like neurological or psychological diseases and whether and how drugs affect and change behaviours in animal models of diseases (Irwin 1968, Teegarden 2012). First, the animal's behaviour is quantified. This is a difficult and complex task, but to understand the complexity of behavioural diversity and its function, it is necessary to catalogue, describe and quantify behaviour. Another important point in the generation of behavioural response is that environment and genes are two substantial and interacting regulators of behaviour. This includes the relationship between genetics and development, the time of performing the testing, the experimenter him- or herself, different characteristics of the animals and the experimental setup (Sousa et al. 2006). Depending on the hypothesis to be investigated, individual test parameters can be altered during diverse test runs. As these variations might include the use of nociceptive impulses, drugs or external stressors, the laboratory mice itself or the experimental environment will be affected. As described above, it will result in behavioural changes, which are used to obtain information about the underlying brain process (Sousa et al. 2006).

4. General Health

The evaluation of general health and physical condition of the animal is important to secure a healthy animal and normal behaviour. It always includes particular observation of the animal to assess noticeable abnormalities or illnesses that could interfere with the different behavioural tests and therefore create artefacts (Irwin 1968, Sousa et al. 2006). According to Crawley (2000) and Sousa et al. (2006), the following aspects of general health should be assessed prior to behavioural testing.

4.1. Abnormal Spontaneous Behaviour

Abnormal behaviour is a behaviour that deviates from normal mouse behaviour as it is described for example by König (König 2012) or Bayes et al. (Bays et al. 2006, König 2012). This includes a wide range of novel behaviours or exaggerated normal behaviour patterns regarding intensity or frequency, like bizzare or excessive comportments, stereotypical behaviours, compulsive licking, self-destructive biting, head flicking, upright walking, retropulsion or spatial disorientation (Hartsock 1982, Irwin 1968, Jensen 2009, Rogers et al. 1997). For the evaluation of this abnormal spontaneous behaviour the mouse is placed in an empty standard mouse cage or viewing jar, then observed for a particular amount of time and quantitated with the help of rating scales and scoresheets (Crawley 2000, Rogers et al. 1997). It has to be taken into consideration that genetic variability in behaviour exists within and between different strains of laboratory mice which, however, only carry a small part of the variation found in wild mice (König 2012). Today, automated computer-assisted test apparatuses are used and video recording is highly recommended (Sousa et al. 2006, Thomas D. Tyler, Richard E. Tessel 1979). Depending on the protocol, different stimulant drugs may be used to induce stereotypic behaviour.

4.2. Activity Rhythms, Sleep and Home Cage Activity Patterns

Observation of home cage activity can be done with the help of several 24 h periods of videotaping and scoring the material by observers (Crawley 2000, Fox 2007, Tang et al. 2002). Evaluation includes for example activity levels, circadian cycles, exploratory activity, food and water intake as well as sleeping patterns (Crawley 2000, Fox 2007, Rogers et al. 1997). Further, sleep and wakefulness can be assessed in detail with the help of implanted electroencephalogram electrodes (Tang and Sanford 2002).

4.3. Appearance in General

This assessment includes the appearance of fur, whiskers and skin, the evaluation of self-grooming and social grooming behaviours, the colouration of ear and footpads as well as the body size of the animals (Crawley 2000, Sousa et al. 2006). Furthermore, it should be paid attention to signs of infections, wounds or fighting behaviour in the home cage (Sousa et al. 2006).

In 2010 a facial expression score for mice was published (Langford et al. 2010) to score facial expression changes for the measurement of pain (Langford et al. 2010, Miller and Leach 2015) and therefore evaluate the severity of burden in mice. Therefore, an observer scores the presence of facial action units, which are facial actions of individual muscles or groups of muscles. The intensity of each facial action unit is coded on a three-point scale and includes orbital tightening, nose bulge, cheek bulge, ear position and whisker change (Langford et al. 2010). The observation and evaluation of these units are done live while observing the animal or retrospectively from still images. However, there might be differences between these methods with live scores being significantly lower (Miller and Leach 2015). It has to be seen that studies indicate that scores of animals which are not in pain are not stringently zero. Furthermore, there is a significant difference in baseline mouse grimace scale scores between males and females and different strains of laboratory mice (Miller and Leach 2015). Additionally, the mouse grimace scale might not be accurate for evaluation in all procedures, since very acute noxious stimuli do not change the score in comparison to baseline (Langford et al. 2010, Miller and Leach 2014). This facial action coding system (FACS) was followed by publications of similar systems in other species (rat: Sotocinal et al. 2011, horse: Dalla Costa et al. 2014, cat: Evangelista et al. 2019, rabbit: Hampshire and Robertson 2015, lamb: Guesgen et al. 2016, sheep: Häger et al. 2017, cattle: Gleerup et al. 2015). Even semi-automated systems for pain assessment using photos of mouse faces have been developed (Ernst, Kopaczka, Schulz, Talbot, Struve et al. 2020, Ernst, Kopaczka, Schulz, Talbot, Zieglowski et al. 2020). But all this pain assessing systems using only the face of an animal neglects other important factors like gait, posture, fur, weight and social interactions which are easily detectible in laboratory mice.

4.4. Body and Skin Temperature

The temperature of the mice is mainly obtained by three different methods. These include the use of a rectal thermistor or telemetry via implantable temperature transponders, which are located subcutaneous or in the abdominal cavity to measure body core temperature and the

use of surface infrared thermometry to measure surface (skin) temperature (Crawley 2000, Mei et al. 2018, Meyer et al. 2017, Saegusa and Tabata 2003). Especially the use of a rectal thermistor might be stressful (Briese and Cabanac 1991) and can lead to injuries and therefore septicemia and death of the animal (Clement 1993). In contrast, telemetry requires anaesthesia and surgery for initial implantation, but body temperature can be measured without any restrainment (Clement et al. 1989, Meyer et al. 2017). Infrared thermometry is a non-invasive method, even though mice might be restrained or kept in a confined space to standardise the exposed surface area and measurement conditions (Meyer et al. 2017). Near-infrared thermography for measuring wellbeing is mainly used in farm animals since-mice pose the problem of a very small surface, influence of coloured fur in different strains (Fink et al. 2015), and body temperature changes of more than one centigrade between active and inactive phases of healthy mice (van Loo et al. 2007). Mice injected peritoneally, which corresponds to mild burden according TVG2012, do not show any difference in body surface temperature (Gjendal et al. 2018). Such differences as indicator for disturbed wellbeing in mice could only be shown in sedated animals (Pereira et al. 2018, Pereira et al. 2019), which is not feasible for phenotyping tests. However, all temperature measurements could be done once or repeated several times.

4.5. Body Weight

Weight of the animals can be obtained by using an animal balance or adapted pan balance. The weight could be obtained once or in daily or weekly intervals (Crawley 2000).

4.6. Neurological Reflexes

Reflexes are important for appropriate behavioural responses, so their evaluation has major importance. The testing implies righting reflexes by placing the animal onto its side or back, the postural reflex by shaking an empty cage sidewards and up and down to watch the postural reflexes of the mouse placed in it; eye blink reflex by approaching a cotton swab towards the eye; ear twitch reflex by touching the ear with a cotton swab; and whisker-orienting reflex by brushing the whiskers of a free moving mouse, watching its reaction (Crawley 2000, Rogers et al. 1997).

4.7. Posture and Gait

This task implies observation of posture and gait during resting and exploration of the home cage or an unknown environment (Crawley 2000, Irwin 1968, Rogers et al. 1997).

4.8. Seizures and Vestibular Dysfunction

Motor abnormalities and seizures are observed, either during the evaluation of general health or during motor functional tests. It includes abnormalities like ataxia, immobility, twitching, waltzing, reeling, tottering, convulsions, erratic movements and running. These observations indicate abnormal electrical discharges, mutations and/or structural defects in brain, spinal motor neurons or vestibular structures (Crawley 1999, Crawley 2000, Rogers et al. 1997).

4.9. Sensory Capabilities

Besides the evaluation of neurological reflexes, the sensory capability is an important point of proper behavioural response (Crawley 1999). Pupil constriction and dilatation as a response to light indicate visual abilities (Crawley 2000, Rogers et al. 1997). Olfactory tests are realized using strong odours while observing eventually occurrent anosmia (Crawley 2000). Another important point is the evaluation of tactile responses on mechanical or thermal stimuli, since the somatosensory system is involved in major operations. This includes perception and reaction to exogenous and endogenous stimuli as well as perception and control of body position and balance (Abraira and Ginty 2013).

5. Classification of Behavioural Tests

The different behavioural tests can be classified into different categories. According to Crawley (2000), the classification could be the following: Motor functions; sensory abilities; learning and memory; feeding and drinking; reproductive behaviours; social behaviours; emotional behaviours including animal models of psychiatric diseases; and reward (Crawley 2000, Sousa et al. 2006).

5.1. Motor Function

Measures of motor function are of high interest, as animals with abnormalities may not be able to fulfil some tasks of behavioural tests (Crawley 2000). Furthermore, misinterpretation regarding higher brain functions is possible, due to general changes in motor function and locomotion (Sousa et al. 2006). The following descriptions include some of the better known but still not all existing procedures, since phenotyping laboratories might use their own developed tests.

5.1.1. Gait Analysis and Locomotor Function

5.1.1.1. Circling

Rotational locomotion or circling can be seen in specific mutant mice, in natural mutant mice with particular defects regarding structures in the ear, cerebellum or spinal motor neurons, might be induced through drugs or could be an indication of unilateral lesions of the nigrostriatal dopaminergic pathway (Crawley 2000, Lee et al. 2001). Circling can be observed and assessed in the home cage or an open field apparatus, e.g. via video tracking, with photocell-based systems or can be measured with the help of commercially available rotometers or rotational behaviour monitoring systems (Antipova et al. 2018, Crawley 2000, Koski et al. 2020, Lee et al. 2001, Pierce and Kalivas 2007). As animals do not experience stress during observation, the test should basically be classified as with “no burden”. If rotational locomotion is induced through drugs or surgery, the observation of rotational behaviour might be classified as “mild” or “moderate”.

5.1.1.2. Cylinder Test

This test is used to evaluate locomotor asymmetry and sensorimotor dysfunction. The mouse is placed in a cylinder-shaped glass with an open-top. The contact of the left and right fore-paws with the cylinder wall during rearing is recorded and assessed (Baskin et al. 2003,

Roome and Vanderluit 2015). Since there are basically no stressful aspects in this procedure, we suggest categorising the test as with “no burden”.

5.1.1.3. Foot Fault Test

This test is designed to assess locomotor function (Starkey et al. 2005). The mice are trained to cross a raised, horizontally oriented ladder or grid. Paw slipping and step failures are recorded and the latency to cross the ladder is measured. The trial is repeated several times, then the average time, as well as the average misplacements, are taken.

Depending on the protocol, different variations of a ladder are used. This includes grids with dimensions of 12 x 57 cm with 1.3 x 1.3 cm diameter openings (Qu et al. 2009), 25.4 x 40.6 cm with 8 x 8 mm diameter openings and a metal wire diameter of 1 mm (Lee et al. 2016), 12 mm square wire mesh with dimensions of 32 x 20 cm (Clarkson et al. 2010), 32 x 20 cm with 12 x 12 mm diameter openings (Li et al. 2015), wire circular grid with diameter of 330 mm and 15 x 15 mm grid squares (Starkey et al. 2005), wire grates with bars of 2.5 mm in diameter, spaced 1 cm apart (Shelton et al. 2008), a horizontal ladder with the rungs spaced 3 cm (Xu et al. 2011) or elevated hexagonal grids of different sizes (Zhang et al. 2002). In all these protocols, the walking platforms are elevated, but only in two protocols the height of 50 cm is described (Clarkson et al. 2010, Li et al. 2015). Some protocols use an electric foot shock for foot faults (Barreto et al. 2010, Xu et al. 2011). The trials are carried out for 2 min (Shelton et al. 2008), 3 min (Lee et al. 2016, Starkey et al. 2005) and 5 min (Clarkson et al. 2010, Qu et al. 2009). Additional three days of training with three trials per day are performed in another protocol (Xu et al. 2011).

Mice have to perform the foot fault test on an elevated ladder or grid. As these animals are known to fear heights (Stowers et al. 2017), a stress factor is given. However, we suggest that the general category of this test should be at most “mild”. Depending on the chosen height, the severity might be different, since a higher elevated apparatus could increase the level of stress. As there is just described one used height of 50 cm, an exact classification regarding the severity is difficult. Further, the classification has to be seen in relation to the openings of the walking surfaces. It has to be taken into consideration that wider openings may possibly lead to a higher rate of foot faults and therefore a higher possibility to fall. Of course, precautions should be done to prevent falling or injuries. However, further investigations need to be done regarding the opening diameter and also the possible impact of the

wire or ladder diameter. We suggest that protocols which use a foot shock should be classified as either with "mild" or "moderate" depending on the intensity of current used. This is because an unescapable nociceptive stimulus will cause stress to the animal. Referring to the impact of foot shocks, important factors are the current density, frequency, intensity and duration of the foot shocks as well as the contact resistance. None of the reviewed protocols mentions all of these factors and hence the impact on the animal's body depends on variation and combination of the particular aspects, further investigation is required, and calculation needs to be done. Lastly, the different protocols do not mention any precautionary measures in case mice might fall down or if injuries were observed. If tests are performed in an unknown environment is left unclear. These findings should be kept in mind considering the evaluation of the severity of different test protocols, since it is likely that an unknown environment might cause stress as well as falling down without any precautionary measures.

5.1.1.4. Footprint Pattern/CatWalk™ XT

This test quantifies gait abnormalities and ataxia. The classical footprint pattern test requires dipping forepaws and/or hindpaws into two different ink colours or photo developer. Then the mouse walks through a dark tunnel covered with white paper. Afterwards, the footprint pattern is analysed by an experimenter (Capoccia et al. 2015, Carter and Shieh 2015, Hamers et al. 2006). Nowadays there are different automated video tracked and computer-assisted tests apparatuses, for example camera assisted open field tests or the CatWalk™ XT system by Noldus (Leroy et al. 2009).

The CatWalk™ XT instrument consists of an enclosed 1.3 m glass walkway, a fluorescent lamp which emits light inside the glass plate, a high-speed colour camera, and a recording and analysis software to assess the recorded gait of the mice (Batka et al. 2014, Hamers et al. 2006, Hetze et al. 2012, Parvathy and Masocha 2013). At the end of the walkway a hiding opportunity, e.g. the home cage (Hetze et al. 2012) or a goal box, can be placed, some food pellets are placed in there optionally (Caballero-Garrido et al. 2017). The apparatus' height above the floor can be set individually and is described in variations of 100 cm (Batka et al. 2014) or 150 cm (Caballero-Garrido et al. 2017). The width can be set individually but is only described with 7 cm (Batka et al. 2014). The mouse must cross the walkway without interruption or turning around. Depending on the protocol used, the crossing has to be completed in minimum three compliant runs with just one second try, in five compliant runs in 15-20 min (Caballero-Garrido et al. 2017) or in three compliant runs in up to maximum 20 overall runs

per day (Batka et al. 2014, Hetze et al. 2012). To accomplish an uninterrupted sequence, descriptions of training define training in just three trials (Hetze et al. 2012) or in two to seven days of training by 25 min of free walking per day (Caballero-Garrido et al. 2017). Training can be accompanied by noises (clapping or rubbing the fingers) to motivate the mouse to go back into the walkway when they come out of this area (Caballero-Garrido et al. 2017). All CatWalk™ XT tests must be done in a dark room.
In contrast, to assess footprint pattern without the CatWalk™ XT instrument, the mouse must be restrained first to cover its paws with non-toxic ink or photo developer (Hamers et al. 2006). The animal is then allowed to walk along a walkway with a sheet of white paper on the floor. The walkway can vary in length and width (Hamers et al. 2006). One protocol is described with 50 cm length and 10 cm width, with 10 cm high walls and an enclosed box at its end (Capoccia et al. 2015). The test is done in an illuminated room.

The testing with CatWalk™ XT may cause stress and anxiety as it is taking part in an unknown and open environment, which leads to anxiety behaviour in mice (Bourin et al. 2007). This anxiety to a novel environment can be reduced by training trials and therefore familiarisation with the apparatus concerning test trials. Proceeding all experiments in a dark environment is also reducing stress, as mice normally prefer an environment without light. These findings lead to the suggestion of categorising the test as at most “mild“. The different protocols do not describe how animals get back into the starting position if a trial is repeated. If it is assumed that mice get put back by hand, the level of stress might increase at first, since handling with a hand is known to cause stress (Hurst and West 2010). With the repetition of handling during test and training trials, a familiarisation effect occurs and stress decreases.
In contrast, footprint pattern requires restraint and handling of the individuum, to cover the paws with ink or photo developer. As already mentioned, restraint and handling are known to cause stress and anxiety (Clarkson et al. 2018, Gouveia and Hurst 2013, 2017, Hurst and West 2010, Pawlyk et al. 2008). Lastly, the test is always described to be conducted in an illuminated room, which also causes anxiety as mice prefer dark environment. We suggest that the classical footprint pattern assessment should be classified at most as “mild”.

5.1.1.5. Rotometer

This test apparatus measures rotational locomotion clockwise and counterclockwise as basically mentioned in 5.1.1.1. The apparatus consists of a transparent bowl-shaped testing area and a transducer. The mouse is then wearing a collar or body harness, which is connected to

the transducer. When the animal is placed into the apparatus, rotations are measured (Pierce and Kalivas 2007). Since animals might experience stress due to the wearing of a body harness or collar, we suggest categorising this test as “mild”.

5.1.1.6. Treadmill Test

Different automatic treadmill systems are used to analyse gait of which the most common one is the DigiGait™ Imaging System. The test animal moves on a motorised transparent treadmill belt with different walking speeds. The mice’s walk is recorded from ventral by a video camera below the treadmill belt. The system automatically vectorizes and pixelates the ventral view and further creates diagrams representing the different parameters of the animal’s walk (Dorman et al. 2014, O. H. Maghsoudi et al. 2015). Some protocols use a shock grid at the end of the lane, to provoke an aversive stimulus (Castro and Kuang 2017, Schefer and Talan 1996). Other protocols use trials until total exhaustion of the mice (Castro and Kuang 2017). Under these circumstances we suggest categorising this test as either “mild”, “moderate” or “severe”, depending on the intensity of current used and if the endpoint is the exhaustion of the mouse. Otherwise we suggest classifying the test as with “no burden”.

5.1.2. Locomotor Activity and Exploratory Behaviour

5.1.2.1. Circadian Wheel Running

To assess the circadian locomotor activity, mice are kept individually in cages equipped with running wheels connected to a device, recording the number of revolutions and speed. The running activity is measured during several days or weeks and different protocols of light/dark circles can be used (Carter and Shieh 2015, Crawley 2000). Additionally, the effects of different drugs, environmental conditions, diets or genetic conditions can be assessed during voluntary wheel running (Sherwin 1998). Since circadian wheel running is voluntary, we suggest categorising the test with “no burden”.

5.1.2.2. Hole Poke Exploratory Test

This test is used to measure the locomotor activity. The apparatus is a box with holes along the walls and/or floor. As mice tend to poke their noses into these holes, Photocell beams measure the head-dipping over a defined period (Crawley 2000). Depending on the used system other parameters could be analysed as well.

The procedure is performed in different types of apparatus. Therefore are described an automatic hole-board apparatus (model ST-1, Muromachi Kikai, Japan), which consists of a grey wooden box (Takeda et al. 1998), a box (Arenas et al. 2014), a four mouse 9-hole chambers (CeNeS, Cambridge, UK) (Trueman et al. 2005) and a 2.2 cm thick board. These apparatus are made up of different materials like grey wood with walls of Plexiglas (Arenas et al. 2014), stainless steel with a roof of Perspex and walls of Plexiglas (Trueman et al. 2005), completely consist of wood (Lalonde 1987), Plexiglas (Flores-Montoya and Sobin 2015) or Perspex grey panels. The various measurements are 50 x 50 x 50 cm (Takeda et al. 1998), 28 x 28 x 20.5 cm (Arenas et al. 2014), 14 x 13.5 x 12.6 cm (Trueman et al. 2005), 70 x 70 cm with 34 cm height of the walls (Lalonde 1987), 50 x 50 cm (Liu et al. 2017), 40 x 40 cm or 40.64 x 40.64 x 40.64 cm (Flores-Montoya and Sobin 2015). The protocols also differ in number and diameter of holes. That is why there are described four holes with a diameter of 3 cm (Takeda et al. 1998), 16 holes with a diameter of 3 cm (Arenas et al. 2014), nine horizontally placed holes with a 1.1 cm diameter (Trueman et al. 2005), 16 holes without mentioned diameter but 2 cm deepness (Lalonde 1987), an elevated plastic floor with 16 holes (Liu et al. 2017) and 16 holes with a 2.54 cm diameter (Flores-Montoya and Sobin 2015). Further, variations of trial durations are presented as the following: 3 min (Flores-Montoya and Sobin 2015), 5 min (Takeda et al. 1998), 10 min (Liu et al. 2017), 5 min with repetition one week later (Labots et al. 2018), trials of 4 min in three consecutive days with two days exploring the apparatus during 10 min in groups of five or six (Lalonde 1987). Also, a protocol with training a serial choice visual discrimination task with trials of 30 min in 23 consecutive days with four days of pretraining procedure is described (Trueman et al. 2005). The apparatus, which is described as a platform, can be positioned 7-18 cm above table level, as it is supported by two brackets. The trials are performed between 10:00-14:00 h (Arenas et al. 2014, Labots et al. 2018) or 08:00-13:00 h (Liu et al. 2017) as described in two protocols. Some protocols mention different reward methods, like offering non-diluted strawberry milk in case of responding right (Trueman et al. 2005).

The hole poke exploratory test takes place in an unknown environment, which is known to cause stress (Bourin et al. 2007, Prut and Belzung 2003, Wilson and Mogil 2001), but can be reduced by familiarisation with the help of training trials. As there is no further impact on the animals, we suggest that the test might be categorised at most as “mild”. It is probable less stressful if the apparatus is made of non-transparent materials, like wood or stainless steel, as mice prefer an enclosed, dark and small space (Crawley 1999). If more holes are used in

the apparatus, the increasing new impressions could have an impact on the animal's stress level, since they have to investigate more holes. Contrary, the impact could also be in a positive way, as more holes could be a contrast to home cage housing and stimulate natural curiosity. Further, studies about the trial length should be done, as they might show how a longer trial length influences the animal's welfare or at which point a familiarisation effect occurs. Studies show that a repeated open field test does not make any difference in stress level for mice (Bodden et al. 2018), therefore repetition of the hole poke exploratory test will not increase the animal's burden in this test. It could be possible that exploring the unknown environment in groups during the pretraining procedure, as described in (Lalonde 1987), reduces stress, but further studies need to be done to evaluate this hypothesis. In addition, investigations might be useful to evaluate if rewarding reduces animal's stress, since they seem to have positive experiences if rewards are used (Spangenberg and Wichman 2018).

5.1.2.3. Open Field Test

The open field test is used to measure fear-related behaviours or to describe and quantitatively evaluate the spontaneous activity, exploratory behaviour and locomotion of rodents as described in the following. An open field arena without any hiding opportunity is used. Shapes and sizes of the apparatus vary throughout protocols and laboratories. Different locomotor parameters can be measured within different time ranges with the help of photocell beams or video tracking systems (Crawley 2000).

The examined protocols use different open field arenas. Described are circular (Fahlstrom et al. 2011) and square (Bodden et al. 2018, Flores-Montoya and Sobin 2015, Liu et al. 2017) shaped arenas with sizes of 50 cm diameter with 30 cm walls (Fahlstrom et al. 2011), 80 x 80 x 42 cm (Bodden et al. 2018), 40.64 x 40.64 x 40.64 cm (Flores-Montoya and Sobin 2015) 45 x 45 cm (Lijam et al. 1997), automated activity monitor system (Versamax, Accuscan, Columbus, OH) with a 50 x 50 cm test chamber (Liu et al. 2017). Exploring times are described with 180 s (Fahlstrom et al. 2011), 5 min (Flores-Montoya and Sobin 2015, Liu et al. 2017), 15 min (Bodden et al. 2018) or 60 min (Lijam et al. 1997). As described in some protocols, the test trials are also repeated two times (Grailhe et al. 1999) or three times (Bodden et al. 2018). Illumination is described with 40 lx (Bodden et al. 2018, Liu et al. 2017). One protocol describes a time of 1 min during which the animal is placed in a cylinder (11 cm diameter and 20 cm height) in one corner of the open field arena at the beginning of the trial, until the cylinder is lifted (Bodden et al. 2018).

The open field test uses the aversive stimuli of an unknown, wide open environment without any hiding opportunity and sometimes bright illumination, which are known to cause anxiety in mice (Bourin et al. 2007, Prut and Belzung 2003, Wilson and Mogil 2001). As there are no physical exhaustion or further manipulation on the animals, we suggest classifying the test at most as either with “no burden” or “mild”, depending on the protocol’s details. A circular open field area has no hiding opportunity in contrast to a square shaped area, where corners could possibly give minimum protection. Indeed, the exact impact of the area’s shape on the animal’s stress level is unclear and needs to be investigated in further studies. Trial time might have an impact on the mice’s welfare, since longer trial times could potentially lead to familiarisation and habituation. Since this might decrease the animal’s stress level, it also has got an impact on the outcome of the procedures. However, it is unclear if habituation is possible within one trial or what trial length would be required to reduce stress and anxiety significantly. Studies indicate that repeated open field testing has no negative impact on animal’s welfare (Bodden et al. 2018).

5.1.3. Motor Coordination and Balance

5.1.3.1. Basket Test

This test is used to evaluate sensorimotor deficits and motor coordination. The mouse is placed in a grid basket, which then is inverted. The time until the animal reaches its home cage by climbing down the mesh walls is measured (Stanford Medicine Behavioral and Functional Neuroscience Laboratory 2020). As there are no stressful aspects in this procedure, the test should be classified as with “no burden”.

5.1.3.2. Beam Walking Test/Balance Beam Test

This test assesses motor coordination and balance (Shelton et al. 2008). While the mouse crosses a horizontal, narrow beam to get into a safe box or a platform, time and missteps are measured.

The balance beam test is described in different variations. The mouse must cross the beam to get into a safe environment, which consists of its home cage, a black box (Almond et al. 2006, Luong et al. 2011, Shelton et al. 2008, Stanley et al. 2005), an enclosed platform (Fan et al. 2010) or just a wider platform. In some cases, the starting point is illuminated brightly to cause an aversive stimulus (Luong et al. 2011, Roemers et al. 2018, Shelton et al. 2008). Respective the beam itself lengths of 60 cm (Almond et al. 2006, Stanley et al. 2005), 75 cm

(Shelton et al. 2008) or 100 cm (Guilford et al. 2011, Luong et al. 2011) are described, as well as a trainings trial beam of 80 cm (Stanley et al. 2005). Normally it is installed horizontally in heights of 20 cm (Almond et al. 2006), 30 cm (Stanley et al. 2005), 50 cm (Luong et al. 2011, Roemers et al. 2018) or 60 cm (Deacon 2013), but a variation of a 30° angled beam is described as well (Roemers et al. 2018). The shape varies between wooden round (Almond et al. 2006, Roemers et al. 2018), square (Roemers et al. 2018) and flat shaped (Luong et al. 2011) beams. Diameters of 5 mm, 6 mm (Roemers et al. 2018), 7 mm, 8 mm (Almond et al. 2006, Stanley et al. 2005), 9 mm (Deacon 2013), 12 mm (Guilford et al. 2011), 15 mm, 22 mm, 28 mm and 35 mm (Deacon 2013) are described. The widths of square and flat beams are described with 4 mm (Roemers et al. 2018), 6 mm (Luong et al. 2011), 12 mm (Luong et al. 2011) and 30 mm (Stanley et al. 2005). In some of the experiments is described that a padded surface (Roemers et al. 2018) or a hammock (Luong et al. 2011) is placed under the beam, to counter injuries, but many experiments do not mention any security for the mice. In some protocols the mouse that fell is positioned again on the beam, to proceed with the trial (Roemers et al. 2018, Stanley et al. 2005). Trainings trials, single trials as well as trials that are repeated three times are described (Guilford et al. 2011, Roemers et al. 2018). One protocol describes various trials using different diameters of rods, ending the test, when the mouse is not able to cross the chosen beam diameter anymore (Deacon 2013). The maximum time of the trials are 60 s (Almond et al. 2006, Stanley et al. 2005) or 120 s (Deacon 2013, Shelton et al. 2008), with resting times of 15 s, 30 s (Roemers et al. 2018), 90 s or a few seconds between the trials and 10 min of resting time between the use of different shaped beams. If the mouse does not move forwards, pushing, poking or prodding by the experimenter are described, to make the animal proceed walking forwards.

Mice have to perform the foot fault test on an elevated beam. As these animals are known to fear heights (Stowers et al. 2017) a stress factor is given. Further, the test takes place in an open spaced environment with a bright light at the beginning of the beam, which causes stress and anxiety in mice (Bourin et al. 2007, Prut and Belzung 2003, Wilson and Mogil 2001). Because of these findings, we suggest categorising the test at most as “mild”. It should be taken into consideration that the burden might increase if test protocols do not use any precautionary measures, like a padded surface or a hammock, since injuries may probably occur by falling, depending on the height of the beam. However, a protocol without any precautions is not recommended at all. It is likely, that walking a longer beam or walking the beam during a longer trial time is more tiring, so the burden possibly increases with the

length of the used beam and an increase of the trial time. Referring to the beam's shape and diameter a flat shaped beam possibly has a positive impact on the mice welfare, since mice probably have a better grip on flat beams. This is why the diameter plays a role as well, as it might be more difficult to hold on a smaller diameter. Obviously, all these findings need to be investigated particularly in additional studies.

5.1.3.3. Rotarod Test

This test measures motor coordination and balance of the tested animals. The mouse is placed onto a rotating roll that gradually increases its speed, so the animal has to move forward continuously. The time the animal can keep running on the roll until it falls off is measured (Crawley 2000, Deacon 2013, Dunham and Miya 1957).

High variability of testing apparatus is given, since there is a great number of manufactures of behavioural testing apparatus. Described are a Smartrod Apparatus (Accusan Instruments, Columbus, OH), which is a single stand-alone unit (Heyser et al. 2013), a stand-alone unit (Med Associates, Inc., St. Albans, VT, USA) (Flores-Montoya and Sobin 2015, Scholz et al. 2015), a MK-660D (Muromachi-Kikai, Tokyo, Japan) (Shiotsuki et al. 2010), an accelerating rotarod (Deacon 2013) by SD Instruments, San Diego, CA (Xu et al. 2011), Type Ugo Basile 7650 (Costa et al. 2001) or Ugo Basile 7600 (Comerio, Italy) (Stanley et al. 2005), at the very beginning of behavioural testing a rolling roller apparatus (Dunham and Miya 1957) and the AccuRotor Rota Rod model RRF/SP (Accusan Instruments, Inc., Colombus, OH) (Bohlen et al. 2009). Heights measure 16.5 cm (Flores-Montoya and Sobin 2015, Scholz et al. 2015), 30 cm (Deacon 2013), 38 cm (Dunham and Miya 1957) and 63 cm (Heyser et al. 2013). As some of the apparatus are not available anymore, height measurements are unreproducible if not mentioned in the protocol. The diameters of the rods are described with 2.75 cm (Xu et al. 2011), 3 cm (Deacon 2013), 3.2 cm (Flores-Montoya and Sobin 2015, Scholz et al. 2015), 3.5 cm (Dunham and Miya 1957), 4 cm (Stanley et al. 2005), 6.3 cm (Heyser et al. 2013) or 9 cm (Shiotsuki et al. 2010). The different rods consist of knurled plastic dowels (Heyser et al. 2013), hard chloroethylene which does not permit gripping on the surface (Shiotsuki et al. 2010), a roller for a window shade (Dunham and Miya 1957) or a surface knurled in a series of parallel ridges along the longitudinal axis, which provides a good grip (Deacon 2013). Some protocols also describe the lane width with 5.7 cm (Flores-Montoya and Sobin 2015) and 122 cm divided into equal compartments with the help of circular cardboard discs (Dunham and Miya 1957). Accelerating speed varies between 5 rpm

(Dunham and Miya 1957), 3.5 rpm to 35 rpm in 300 s (Flores-Montoya and Sobin 2015), 4 rpm to 40 rpm in 300 s (Costa et al. 2001, 2001, Scholz et al. 2015), 4 rpm to 20 rpm with a maximum speed of 40 rpm (Deacon 2013), 5 rpm to 10 rpm in 300 s (Xu et al. 2011), 20 rpm in 300 s with a maximum speed of 99.9 rpm (Bohlen et al. 2009) or no acceleration but a constant speed at 16 rpm (Stanley et al. 2005). The reviewed protocols describe different procedures which include a trial at a constant speed of 16 rpm in three consecutive trials with maximum 120 s and another trial after drug administration (Stanley et al. 2005), four trials of 300 s with a pause of 3 min (Flores-Montoya and Sobin 2015), a pretraining of nine days with increasing rotation speed from 5 rpm to 10 rpm over 5 min with a maximal duration of 300 s, after that three trials per day on two days and further 1-4 weeks after surgery (Xu et al. 2011), a trial with habituation time of 3 min, then a habituation time of 1 min every day before the trial. The trials have an acceleration of 10 rpm and mice are placed back on the rotarod immediately after they fall down, up to five times. The trials are performed once per day in four consecutive days (Shiotsuki et al. 2010). Four mice are tested simultaneously on the rotarod ten times a day, four days in a row (Scholz et al. 2015).

Animals have to perform the rotarod test on an elevated rod. As these animals are known to fear heights (Stowers et al. 2017) a stress factor is given. Further, the test takes place in an open spaced environment, which causes stress and anxiety in mice (Bourin et al. 2007, Prut and Belzung 2003, Wilson and Mogil 2001). However, we suggest categorising the test as at most “mild”. Depending on the height used the burden might differ. Test protocols without any precautionary measures, like a padded surface or a hammock, might have a higher burden, as injuries may probably be caused by falling down according to the hight used. However, a protocol without any precautions is not recommended at all. Besides, a longer trial time, as well as the amount of trial repetition, might possibly be more tiring for the animals. That is why the burden possibly increases in correlation with trial length and number of trial repetitions. Further, the accelerating speed may also have influence, since a faster acceleration time to maximum speed may likely cause the animal to fall easier. The burden might therefore be changing with the maximum speed used during the procedure. A bigger diameter of the rod might eventually decrease stress because animals have a bigger surface for walking and better grip on the rod. Clearly, these findings could be investigated particularly in additional studies for specific evaluation and impact on the animal’s welfare.

5.1.3.4. Staircase Test

This test evaluates motor-fine coordination. The apparatus consists of a start chamber, a transparent narrow corridor, in which the animal cannot turn around. A plinth is located centrally in the corridor and surrounded by a staircase with eight steps on both sides. Food pellets can be placed into a small deepening in the staircase. The mouse climbs on the plinth and has space to reach the reward on each stair, on the left side with the left paw and on the right side with the right paw. The number of rewards taken without error is counted and evaluated in combination with the difficulty (lowest step: easy, highest step: very difficult) (Baird et al. 2001). As this procedure is performed voluntarily by the animal, the test should be classified as with “no burden”. However, if animals are food deprivated the severity might change.

5.1.3.5. Vertical Pole Test

This test measures motor coordination and balance of the animal. The mouse is placed in the centre of a pole which is held horizontally. Depending on the protocol length and diameter of the pole could vary. First, the pole is held horizontally and then slowly inclined up to a 45° position. The animal may walk up or down the pole, to not fall off. Latency time and angle until the mouse falls off the pole are measured (Carter and Shieh 2015, Crawley 2000). As there are no physical exhaustion or further manipulation on the animals, we suggest classifying the test at most as “mild”.

5.1.4. Muscular Strength

5.1.4.1. Grip Strength

This test is used to determine muscle strength. The mouse clings its front paws, hind paws or all paws on a commercially available force meter. Then the animal is held by the tail and slowly and carefully pulled back until the grasp is broken and the animal leaves the grid. The experiment is repeated a set number of times after a recovery break and the highest value is recorded (Boissier and Simon 1960, Luca 2008).

The grip strength test is normally used to measure forelimb strength, but variations are described measuring all four limb or just hind limb strength. The different apparatuses are equipped with different shaped grasps, described as t-shaped metal bars (Model 47200, Ugo-Basile, Varese, Italy) (Montilla-Garcia et al. 2017), rings (Cabe et al. 1978), grids (Columbus instruments) (Roemers et al. 2018), wire bars with 1 mm of diameter and 3 cm gripping surface using the Mark-10 digital force gauge (model M3-025; Mark-10 Corporation)

(Alamri et al. 2018), spring weigh scales (Fisher Scientific, Tustin, CA) with an attached trapeze-shaped bar (Almond et al. 2006), brass triangles (Chatillon Model, Greensboro, NC) (Nevins et al. 1993) or metallic wires (Boissier and Simon 1960). In most cases, the mouse is lifted by the base of its tail (Nevins et al. 1993, Alamri et al. 2018, Montilla-Garcia et al. 2017, Roemers et al. 2018) and held towards the apparatus, so it can grasp the handle. In one experiment it is described that the mouse is lifted about midway along the length of the tail and the body weight is supported by the other hand (Cabe et al. 1978). Then the mouse is gently pulled away horizontally from the apparatus (Cabe et al. 1978, Montilla-Garcia et al. 2017, Nevins et al. 1993, Roemers et al. 2018) until the grip is lost. Variations, pulling the mouse vertically up (Alamri et al. 2018) or down (Almond et al. 2006), are described. A different variation describes the testing of just one forepaw at the time, by covering the other paw with tape, so it cannot be used by the mouse (Alamri et al. 2018). Another is describing the testing of hind paw strength, where the animals first have to grasp a wire mesh with their forepaws, then grasping the grip strength apparatus with their hind paws (Montilla-Garcia et al. 2017). Releasing the mouse to hang on a thin thread or metallic wire with its forepaws, at about 30 cm height, so it will catch the wire with its hind paws to climb up, is another variation of the grip strength test (Boissier and Simon 1960). Repetition of the test is described from three times forelimb (Almond et al. 2006, Cabe et al. 1978, Nevins et al. 1993), three times forelimb and three times four limb (Sheth et al. 2018), four times forelimb and four times four limb test (Roemers et al. 2018), five times (Alamri et al. 2018), six times (Montilla-Garcia et al. 2017) and in one case 15 times. In between the trials no resting time, a few seconds (Alamri et al. 2018), 60 s (Roemers et al. 2018) or 3 min (Nevins et al. 1993) of resting time are described.

The animal is pulled horizontally or vertically up or down and a variable number of trial repetitions is conducted. As mice possibly experience stress due to handling but no further fear or pain during this test, we suggest categorising it as either with “no burden” or “mild”. A higher number of repetitions may cause exhaustion and therefore increase the burden. However, a longer inter trial time could probably have a positive impact on stress and exhaustion, because animals have some time for physical recovery. It is necessary to note that further investigations about the impact of trial repetition and inter trial time on exhaustion of muscle strength in animals are required. Additionally, it should be taken into consideration, that testing just one paw at a time is likely to be more exhausting. Lastly, the process of taping the

unused paw may cause even more stress, as additional restrainment might be required (Clarkson et al. 2018, Gouveia and Hurst 2013, 2017).

5.1.4.2. Wire Hang Test

The test detects neuromuscular abnormalities measuring balance, grip strength and the prehensile reflex (Carter and Shieh 2015, Crawley 2000, Shukitt–Hale et al. 1999). The mouse is placed on a grid, the grid is shaken to make the animal grab the grid and then turned over and hung upside down (Carter and Shieh 2015, Crawley 2000). When placing the mouse on a wire, it is placed in a hanging position directly. The latency time until the animal falls is measured (Carter and Shieh 2015, Crawley 2000).

The different protocols describe various types of hanging wire apparatus. One with a grid of 12 x 12 cm with 7.5 cm high walls (Sterniczuk et al. 2010), one with a grid of 18 x 18 cm, a diameter of 1 mm, spaced 1 cm apart (Lijam et al. 1997), a wire of 2 mm diameter (Shukitt–Hale et al. 1999), a wire grid with 2 mm diameter, surrounded by a small wooden frame (Roemers et al. 2018), a horizontal wire with diameter of 1.5 mm (Gerlai et al. 2000), a standard linear wire hang apparatus consisting of a plastic box measuring 55 x 40 x 35 cm with a wire diameter of 2.5 mm (Hoffman and Winder 2016), a modified wire hanging apparatus with a wire of 2.5 mm in diameter arranged circularly with a diameter of 32 cm, suspended on a pulley in a plastic box of 55 x 40 x 35 cm (Hoffman and Winder 2016), a 55 cm wide metallic wire with a diameter of 2 mm (either a plain wire or a multi-stranded twisted wire, possibly plastic-coated) secured to two vertical stands (Aartsma-Rus and van Putten 2014, van Putten 2011), a wire lid of a conventional housing cage (Oliván et al.), a metal cloth hanger with 2 mm of diameter taped to a shelf (Aartsma-Rus and van Putten 2014) or last a handmade square or the lid of a big cage for a rat or rabbit (Aartsma-Rus and van Putten 2014). The height described variated between 20 cm (Gerlai et al. 2000), 25 cm (Aartsma-Rus and van Putten 2014), 35 cm (van Putten 2011), 37 cm (Aartsma-Rus and van Putten 2014), 50 cm (Roemers et al. 2018) and 62 cm (Shukitt–Hale et al. 1999). There are protocols which describe a padded surface underneath the wire or grid to avoid injuries in case of falling (Aartsma-Rus and van Putten 2014, Roemers et al. 2018, van Putten 2011) and some protocols which mention a hard surface (Gerlai et al. 2000, Shukitt–Hale et al. 1999). Further, the different proceedings test forelimbs (Shukitt–Hale et al. 1999, van Putten 2011) and fourlimbs (Aartsma-Rus and van Putten 2014, Hoffman and Winder 2016, Lijam et al. 1997, Oliván et al., Roemers et al. 2018, Sterniczuk et al. 2010, van Putten 2011). The trial

time variates between 60 s (Lijam et al. 1997), 180 s (Hoffman and Winder 2016, Oliván et al., van Putten 2011) 5 min (Gerlai et al. 2000), 10 min (Aartsma-Rus and van Putten 2014, van Putten 2011), 20 min (Roemers et al. 2018) or no limit (Hoffman and Winder 2016). Trials are performed once (Lijam et al. 1997, van Putten 2011) or three times (Hoffman and Winder 2016, Oliván et al., Sterniczuk et al. 2010, van Putten 2011). The animal is put on the wire again once (Gerlai et al. 2000, Roemers et al. 2018) up to two times (Aartsma-Rus and van Putten 2014) or until the trial time is reached if the mouse falls down or reaches the end of the wire (Hoffman and Winder 2016). Resting time between trials are described with 30 s (Hoffman and Winder 2016) or 5 min (Sterniczuk et al. 2010). Mice were either handled by their tail (Aartsma-Rus and van Putten 2014, Gerlai et al. 2000, van Putten 2011) or via a non-aversive method like described by Hurst and West (2010) (Hoffman and Winder 2016).

We suggest that the wire hang test should be classified at most as “mild”. Furthermore, the test is performed on an elevated wire or lifted wiremesh and hence data suggests that these animals fear heights (Stowers et al. 2017), another stress factor is given. As already described in previous tests the burden might change depending on the height used. Although observations indicate that falling off 20 cm height on a hard surface does not lead to any physical injuries (Gerlai et al. 2000), a protocol without any precautions is not recommended at all and should therefore be classified with higher severity. In addition, a longer trial time, as well as the number of trial repetitions, might possibly be more stressful for the animals as they probably could be exhausted. That is why burdens possibly increase in correlation with trial length and number of trial repetitions. Additionally, a longer inter trial time might have a positive impact on stress and exhaustion because animals have time for physical recovery. Further investigations about the impact of trial repetition and inter trial time on exhaustion of muscle strength in animals are required. It is likely that there are differences between protocols where the animal uses its forelimbs and protocols where it uses all four limbs because using all limbs might be easier and therefore less stressful. Further, there might be a difference in animals that can climb onto and then balance upon the wire without being constantly in an upside down position, as this position could lead to faster exhaustion.

5.2. Sensory Abilities

Sensory capabilities are important for proper behavioural responses, as already mentioned in 4.9. abnormalities may indicate mutations in the affected corporal area. In the following, different tests are described to evaluate the animal’s sensory abilities.

5.2.1. Olfactory Acuity and Olfactory Discrimination

Simple tests are used to evaluate olfactory acuity. This might be through a hidden, good smelling object measuring the time until the mouse locates the object. Another possibility is to quantitate the time the animal spends sniffing an attractive, unknown odour distributed on the cage. The time sniffing a neutral odour, for example a drop of water, is used as control (Yang and Crawley 2009). The olfactory discrimination test for the evaluation of olfactory acuity is described in the following.

In this test, the mouse must choose between two odours to obtain a food reward. The apparatus can be shaped in different ways, but basically distinct odours are delivered by an air flush and the animal has to differentiate between the odours for being rewarded (Crawley 2000). However, varying protocols are described.

There are different types of olfactory discrimination protocols. In the reviewed ones the test takes place in the animals' home cage (Enwere et al. 2004, Huang et al. 2019, Siopi et al. 2016, Tillerson et al. 2006), a computer controlled olfactometer (Siopi et al. 2016), a Y-maze (Bowers and Alexander 1967) a go/no-go box (Kunkhyen et al. 2018) or computer assisted two hole board apparatus (Mandairon et al. 2006). Different odours and odorant mixtures with different compounds are used. This includes cinnamon (Bowers and Alexander 1967, Tillerson et al. 2006), paprika (Tillerson et al. 2006), amyl-acetate, peppermint (Kunkhyen et al. 2018) isoamyl-acetate, N-butanol, carvone enantiomeres (Siopi et al. 2016), oil of juniper (Bowers and Alexander 1967), odours of different male and female mice in form of bedding or urine (Bowers and Alexander 1967, Kunkhyen et al. 2018, Tillerson et al. 2006), coconut extract, almond extract (Enwere et al. 2004), limonene (Huang et al. 2019, Mandairon et al. 2006), propionic acid (Mandairon et al. 2006), 2-heptanol (Huang et al. 2019), which are diluted in water (Enwere et al. 2004, Kunkhyen et al. 2018, Siopi et al. 2016, Tillerson et al. 2006), denatonium benzoate (Enwere et al. 2004) or mineral oil (Huang et al. 2019, Siopi et al. 2016). These odours are presented on a glass plate, wooden blocks (Tillerson et al. 2006), through the odour port of a flow olfactometer (Bowers and Alexander 1967, Kunkhyen et al. 2018, Siopi et al. 2016), in a culture dish (Enwere et al. 2004), glass plates with filter papers (Huang et al. 2019) or on a polypropylene swab (Mandairon et al. 2006). The protocols describe six sessions of four trials per day, five trials (Mandairon et al. 2006), six trials (Enwere et al. 2004, Huang et al. 2019, Tillerson et al. 2006), eight blocks á 20 trials per day (Siopi et al. 2016), 20 trials per day with a maximum of 140 trials in total if no learning was

documented (Bowers and Alexander 1967), 100 trials per day in four consecutive days (Kunkhyen et al. 2018). Trial time is described with 2 min (Enwere et al. 2004, Tillerson et al. 2006) or 3 min (Mandairon et al. 2006, Tillerson et al. 2006), The inter trial time varies between 30 s (Enwere et al. 2004) and 15 min (Mandairon et al. 2006, Tillerson et al. 2006). Most protocols work with training and habituation trials (Enwere et al. 2004, Huang et al. 2019, Kunkhyen et al. 2018, Mandairon et al. 2006, Siopi et al. 2016, Tillerson et al. 2006). Also, some protocols use food or water deprivation either during training sessions or during trials (Bowers and Alexander 1967, Enwere et al. 2004, Siopi et al. 2016) while others use a reward (Mandairon et al. 2006).

The olfactory discrimination test itself has no negative impact on the animals, which is why it could be categorised as either with “no burden” or “mild” depending on the following protocol details. Although compounds like amyl-acetate, isoamyl-acetate, n-butanol, propionic acid, 2-heptanol, denatonium benzoate or mineral oil are naturally occurring compounds, they could be, depending on the concentration, noxious (Institut für Arbeitsschutz der Deutschen Gesetzlichen Unfallversicherung 04.09.2020, Sabbath et al. 2014). Increased exposure to these compounds might possibly induce a hazardousness to the mice. Another major factor is deprivation of food and/or water, as data suggests that both have a high negative impact on the animals, which is why it can lead down to malformation in pregnant mice (Kalter 1954, Rogoyski 1966, Rosenzweig and Blaustein 1970, Runner, M. N., Miller J. R. 1956). However, there might be a positive impact on mice if rewards are used (Spangenberg and Wichman 2018).

5.2.2. Visual Acuity

Visual acuity might be evaluated with the help of simple reflex responses which have been mentioned in 4.1.6. and 4.1.8. Additionally, the following tests are used for further evaluation.

5.2.2.1. Light-Dark Test

This test evaluates two main parameters. One is the ability to distinguish light from dark, the other is evaluation of anxiety. The apparatus consists of a bigger open and light area and a smaller covered and dark area, which are connected with an opening in between the two compartments. It is based on the fact that mice naturally avoid bright areas without coverage and therefore prefer to stay in covered, darker areas (Bourin and Hascoet 2003, Crawley and Goodwin 1980).

The different apparatus are described as rectangular boxes with different measurements of 46 x 27 x 30 cm (Hascoet and Bourin 1998), 42 x 21 x 25 cm (Serchov et al. 2016), 45 x 27 x 27 cm (Heredia et al. 2014, Martin and Brown 2010), 46 x 20 x 30 cm (Colla et al. 2015), 56 x 33 x 30 cm (Ardayfio and Kim 2006), 40 x 42 x 26 cm (Gapp et al. 2014) and 48 x 24 x 24 cm (Paiva et al. 2010). In one case an open field area of 30 x 30 x 20 cm was used (Kulesskaya and Voikar 2014). The apparatus is divided into a larger light compartment and a smaller dark compartment in all reviewed papers except one which divides the open field area into two equal areas (Kulesskaya and Voikar 2014). The light compartment is painted white (Hascoet and Bourin 1998, Heredia et al. 2014), the floor, walls and lid are made out of clear Plexiglas (Martin and Brown 2010), out of white Plexiglas (Paiva et al. 2010) or white plastic (Gapp et al. 2014). The light compartment is described with measures of 27 x 27 cm (Hascoet and Bourin 1998, Martin and Brown 2010), 2/3 of the whole apparatus (Ardayfio and Kim 2006, Colla et al. 2015, Gapp et al. 2014, Heredia et al. 2014, Serchov et al. 2016) or 15 x 15 cm (Kulesskaya and Voikar 2014). One protocol also mentions that the light compartment is larger, but measures light and dark compartment with 24 cm length and 24 cm width, both (Paiva et al. 2010). Additionally, the compartment is illuminated in different ways. Described are a really bright light with 60 W and 400 lx (Hascoet and Bourin 1998), an illumination of 130 lx (Gapp et al. 2014), 200-400 lx (Serchov et al. 2016), 350 lx (Heredia et al. 2014), 640 lx by a 60 W bulb placed 40 cm above the centre of the compartment (Martin and Brown 2010), illumination by a 60 W lamp (Ardayfio and Kim 2006, Paiva et al. 2010) or by two 40 W light bulbs with 1000 lx which are located 50 cm above the floor (Kulesskaya and Voikar 2014). The dark compartment is painted black (Hascoet and Bourin 1998, Heredia et al. 2014), made out of black Plexiglas (Martin and Brown 2010, Paiva et al. 2010), black plastic walls and lid (Gapp et al. 2014) or it consists of an insert with black walls and lid which is non-transparent for visible light (Kulesskaya and Voikar 2014). Therefore it is described with measurements of 18 x 27 cm (Hascoet and Bourin 1998, Martin and Brown 2010), 1/3 of the whole apparatus (Ardayfio and Kim 2006, Colla et al. 2015, Gapp et al. 2014, Heredia et al. 2014, Serchov et al. 2016) or 15 x15 cm (Kulesskaya and Voikar 2014). For illumination are described a dim red light with 60 W and 4 lx (Hascoet and Bourin 1998), illumination with 5 lx or less (Serchov et al. 2016), but no information about illumination was given in the other reviewed papers (Ardayfio and Kim 2006, Colla et al. 2015, Gapp et al. 2014, Heredia et al. 2014, Kulesskaya and Voikar 2014, Martin and Brown 2010, Paiva et al. 2010). To separate light and dark compartment, a partition with different

sized openings is part of the apparatus. Sizes are described with 7.5 x 7.5 cm (Hascoet and Bourin 1998, Heredia et al. 2014, Martin and Brown 2010), 3 x 4 cm (Serchov et al. 2016), 7 x 5.5 cm (Kulesskaya and Voikar 2014), 5 x 5 cm (Gapp et al. 2014), 8 x 8 cm (Paiva et al. 2010) or a vertical sliding door of 8 cm (Ardayfio and Kim 2006). In the protocols, the animals are put in the middle of the light area (Colla et al. 2015, Gapp et al. 2014, Heredia et al. 2014, Paiva et al. 2010, Serchov et al. 2016) facing away from the opening (Hascoet and Bourin 1998) or facing the opening (Martin and Brown 2010) or put in a corner facing away from the opening (Kulesskaya and Voikar 2014). Another protocol describes animals placed in the dark compartment facing the light compartment (Ardayfio and Kim 2006). The trial time varies between 5 min (Colla et al. 2015, Hascoet and Bourin 1998, Heredia et al. 2014, Martin and Brown 2010, Paiva et al. 2010, Serchov et al. 2016) and 10 min (Ardayfio and Kim 2006, Gapp et al. 2014, Kulesskaya and Voikar 2014). Additionally, one protocol describes the use of an apparatus contaminated with the dirt of other mice to reduce neophobic responses (Hascoet and Bourin 1998).

The Light-Dark test uses aversive stimuli like a bright illumination and white or bright coloured walls in an unknown, open environment. Further, the dark compartment, which is closed and smaller, less illuminated and coloured dark, is also unknown but not paired with other aversive stimuli. No harmful or invasive stimuli are used, and animals have the opportunity to escape the light compartment by entering the dark one. Therefore, this test is not connected with any burden for the animals and should be categorised as with “no burden”. Placing the animals in the light department might induce a little stress, but they can avoid it by changing to the dark compartment.

5.2.2.2. Visual Cliff Assay

This test investigates depth perception and gross visual ability (Fox 1965). It consists of a box with two levels, an upper and a lower one. They are connected through a vertical drop which is approximately 0.5 m deep. A chequerboard pattern is placed on both planes, but for safety reasons a Plexiglas is placed over the drop, extending the upper plane and creating a visual cliff. The mouse is placed onto the centre ridge then latency time is measured until the animal moves off the ridge onto the safe chequerboard pattern zone. Blind mice will move forward across the cliff without pause (Fox 1965). As far as mice will experience no fear or pain the test should be regarded as with “no burden”.

5.2.2.3. Visual Water Box

This test assesses the visual function in combination with a morris water maze. Therefore, a trapezoidal-shaped pool with two computer-controlled monitors at one end is used. Midline dividers extend from in between the monitors. A hidden platform is placed underneath one of the monitors. To escape the water, mice have to distinguish a low spatial frequency sinewave grating from homogeneous grey. Pretraining, task training and acuity testing are the three phases of this task (Prusky et al. 2000). We suggest categorising the visual water box test as either “mild” or “moderate” depending on the trial time.

5.2.3. Auditory Acuity and Acoustic Startle Response

To assess the auditory acuity in mice, different approaches can be used. One is the use of electrophysiological measures, which will not be described further, since it is no behavioural assay (Willott 2001). Further tests might be the observation of pinna movements, evaluation of auditory acuity with the help of avoidance conditioning or the approach the source of a sound (Willott 2001). The acoustic startle response as a behavioural test will be described in the following.

This test measures the gross hearing ability and auditory threshold. Loud sounds (more than 100 dB) cause motor reflexes in rodents, which are considered to be protective mechanisms. Brief sounds (75-120 dB) are delivered into a small cylinder, where the mouse is placed. A white background noise (level 70-75 dB) is present. The flinch amplitude of the animal is measured by an automated system. The acoustic startle threshold is the minimum decibel level that causes flinches. Various brain regions, which are also important in emotion control, are involved in these startle responses (Crawley 2000, Valsamis and Schmid 2011), which are often experienced as aversive stimuli. We suggest categorising the acoustic startle response as either “mild” or “moderate”.

5.2.4. Taste Acuity

This sense is generally measured in choice tests using food or water. The most generally used test is the two-bottle choice test. Two similar bottles with different taste solutions are placed in the cage and the consumed volume is measured, so that preferences can be seen. To avoid place preferences during the test time, their location is switched randomly. To evaluate taste acuity, different concentrations of the solution are used (Carter and Shieh 2015,

Crawley 2000). As far as mice will experience no fear or pain the test should be regarded as with “no burden”. The taste acuity is also used for associative learning tests (see 5.3.2).

5.2.5. Tactile Acuity

To assess the animal’s tactile acuity, simple reflex responses to tactile stimuli, the air puff flinch response, the T/Y-Maze and the Von Frey test are used. These are described in the following.

5.2.5.1. Air Puff Flinch Response

In this test, the animal is disturbed by puffs of compressed air. It is similar to the acoustic startle response described in 5.2.3. but as a tactile version (Crawley 2000, Wahlsten 2011). Therefore, we suggest categorising the test as either “mild” or “moderate”.

5.2.5.2. T/Y-Maze

In this test, the mouse learns to find a reward on the right or left arm of a T/Y-shaped maze. As the animal is motivated by food, a feed withdrawal is done during test preparation. To get information about touch sensitivity on the paws, the floor is covered with different textures. Discrimination between the arms is then measured (Crawley 2000). As far as mice will experience no fear or pain the test should be regarded as with “no burden”.

5.2.5.3. Von Frey Test

The Von Frey Test is used to measure the sensitivity of touch by using the mechanical stimuli of calibrated Von Frey filaments, determining the 50% pull-off threshold. The filaments are inserted through the holes of a close-meshed wireframe platform, touching the plantar surfaces of the animal’s paws that are collocated on top. The paw is pulled away when reaching the reaction threshold (Deuis et al. 2017, Santos-Nogueira et al. 2012). Chaplan et al. (1994) describe withdrawal of the paw to stimuli of lower intensity after a particular number of repetitions. Hence the animal is located in a narrow Von Frey apparatus, we suggest categorising the test as at most “mild”.

5.2.6. Nociceptive Sensitivity

Tests to evaluate the animal’s nociceptive sensitivity are the flinch test, formalin test, hot plate test, plantar (Hargreaves) test, Randall-Selitto test, tail flick test or the writhing test that are described in the following.

5.2.6.1. Flinch Test

This test assesses sensitivity to electrical irritation. The mouse is placed into a chamber with an electrifiable grid floor. Consecutive trials of increasing intensity in electric foot shocks are exerted until twitching, vocalisation or flinching is observed (Blake et al. 1963). As animals experience pain and fear, we suggest categorising the test as "moderate".

5.2.6.2. Formalin Test

This test evaluates the acute pain response to an injected noxious chemical. A formalin solution is injected subcutaneously into the mouse's hind paw, then pain-related responses are recorded. The test is divided into two response phases. The first phase, also called the early phase, is an acute peripheral pain response from pain fibre activity and begins immediately after injection up to 10 min. The second phase, also called the late phase, is the central nociceptive sensitization due to tissue damage. It begins about 5-15 min after injection and is recorded up to one hour. The two phases are divided through a quiescent period (Bannon and Malmberg 2007, Carter and Shieh 2015, López-Cano et al. 2017)

First, the different test protocols use different formalin solutions of 1 % (Hunskaar et al. 1985), 2 % (Carey et al. 2017, O'Brien et al. 2013), 2.5% (Nikoui et al. 2016) or 5 % (Bannon and Malmberg 2007, Hunskaar et al. 1985, Nakamoto et al. 2010). Then, the different ways of injection are described with intraplantar subcutaneous injection (Bannon and Malmberg 2007, O'Brien et al. 2013), injection into the dorsum of the paw (Masocha et al. 2016), the superficial surface (Carey et al. 2017), the subplantar space (Nakamoto et al. 2010) or intradorsal subcutaneous injection (Hunskaar et al. 1985, Nikoui et al. 2016). Injection volume varies between 10 µl (Carey et al. 2017, Nakamoto et al. 2010, O'Brien et al. 2013) and 20 µl (Hunskaar et al. 1985, Nikoui et al. 2016). The used apparatus for testing might be a plexiglass platform in behavioural chambers (O'Brien et al. 2013), a cylindrical test chamber with a metal band placed around the animal's paw (Masocha et al. 2016), a clear Plexiglas chamber with an elevated platform (Carey et al. 2017), a standard macrolone cage of 30 x 12 x 13 cm with a mirror placed behind the chamber (Hunskaar et al. 1985), a funnel where the mouse is placed under onto a surface, both made of glass and with a mirror placed underneath (Nikoui et al. 2016) or a clear container with mirrors behind and beside (Bannon and Malmberg 2007). Acclimatisation time is described with 15 min (Carey et al. 2017), 15-30 min (Bannon and Malmberg 2007), 1 h (Masocha et al. 2016), 2 h (O'Brien et al. 2013) or 18 h to the environment and 2 h to the apparatus (Hunskaar et al. 1985). The

described observation time was described with 30 min (Nakamoto et al. 2010, Nikoui et al. 2016), 40 min (Bannon and Malmberg 2007, Masocha et al. 2016) and 60 min (Carey et al. 2017, Hunskaar et al. 1985, O'Brien et al. 2013). Furthermore, one protocol describes that the animal's injected paw should be marked with a permanent black marker and for this purpose be restrained by one person and marked by another or put into a restrainer (Bannon and Malmberg 2007). Only two protocols mention the equipment for injection, which is a micro syringe with a 26-gauge needle (Bannon and Malmberg 2007, Hunskaar et al. 1985) or a 50 µl Hamilton syringe with a 30-gauge needle (Bannon and Malmberg 2007).

The formalin test is painful, and tissue damaging and should therefore be classified as either "mild" or "moderate", depending on the level of pain and tissue damage. Only concentrations below 0.2 % lead to a short-lasting response with few histological tissue changes (Rosland et al. 1990). However, there are protocols just mentioning the equipment used for the injection. Furthermore, the animal needs to be restrained either by hand or by a restrainer, which are both stressful methods (Clarkson et al. 2018, Gouveia and Hurst 2013, 2017, Hurst and West 2010, Pawlyk et al. 2008, Rosenzweig and Blaustein 1970). Therefore, it has to be seen, that a shorter duration of formalin application is less stressful for the animals. Additionally, the safety of the application increases with the diameter of the injection needle regarding breaking or bending. A metal band around the paw, as described in Masocha et al. (2016), could increase the animal's stress as it is kind of a long-lasting restraint as well.

5.2.6.3. Hot Plate Test

The test animal is placed on a heating plate surrounded by a plastic cylinder. The latency time until showing an obvious nociceptive response, such as licking of a hind paw, vocalization, hopping or escape response is measured (Bohn et al. 2002, Deuis et al. 2017, Woolfe and Macdonald 1944).

The hot plate test is normally carried out with plate temperatures of 55±1 °C (Bannon and Malmberg 2007, Masocha et al. 2016, Menéndez et al. 2002, Minett et al. 2014), but there are also test protocols describing temperatures of 42 °C (Tjolsen et al. 1991), 50 °C (Bohn et al. 2002, Minett et al. 2014), 55 °C (Bannon and Malmberg 2007, Minett et al. 2014), 53 °C, 56 °C (Bohn et al. 2002) or 55 °C up to 70 °C in steps of 5 °C (Woolfe and Macdonald 1944). Modifications of the hot plate test, with steadily increasing temperature (Tjolsen et al. 1991) or just testing one hind paw at the time (Menéndez et al. 2002), are described as well. The

starting temperature in the increasing temperature hot plate apparatus is 30 °C or 42 °C (Tjolsen et al. 1991), using different heating rates per minute until a response of the animal is observed. To conduct handling of the mice two possibilities are mentioned. Restraint by the hand of the experimenter (Menéndez et al. 2002) or just placing the mouse into the apparatus without restraint (Minett et al. 2014, Tjolsen et al. 1991). One of the protocols therefore describes the apparatus two plates surrounded by a Plexiglas chamber and one plate kept at a constant temperature while the other was set to test temperatures (Minett et al. 2014). Cut-off times of 20 s (Masocha et al. 2016), 30 s (Bannon and Malmberg 2007, Bohn et al. 2002, Menéndez et al. 2002, Woolfe and Macdonald 1944) or 180 s (Bannon and Malmberg 2007) are set. Testing intervals of 2 min (Menéndez et al. 2002, Minett et al. 2014), 5 min (Tjolsen et al. 1991), 10 min in the first hour and 20 min in the second hour (Woolfe and Macdonald 1944a) or an unlimited time are described, depending on the drug which is tested. Nociceptive responses are observed and measured manually by the experimenter observing the animals (Bohn et al. 2002, Menéndez et al. 2002, Tjolsen et al. 1991) or via video tracking systems (Minett et al. 2014).

The hot plate test causes a nociceptive response, which is inescapable for the mouse in most of the protocols. It also takes part in an open spaced and novel environment, which causes mice to exhibit fear and show anxiety behaviour (Bourin et al. 2007, Wilson and Mogil 2001). Nociceptive responses are evaluated by observation. This is why the evaluation could be subjective or misinterpreted, depending on the experimenter and its experience. Licking of front paws, for example, is normally not used as a nociceptive response, as it can be seen very commonly as grooming behaviour. However, that does not mean that it cannot be a sign of a nociceptive response as well. It should be taken into consideration that the experimenter’s latency time until noticing the mouse’s response, processing the information and turning off the plate could be subject to variabilities. Because of these findings, we suggest categorising the test as “moderate“. There is some evidence that restraint by hand might be a more stressful procedure, as it is known that restraint and handling both cause stress and anxiety in mice (Clarkson et al. 2018, Gouveia and Hurst 2013, 2017, Hurst and West 2010, Pawlyk et al. 2008). Immobilisation by hand could also be considered as positive because the experimenter just focuses on one or two paws at a time and therefore it is more likely that a nociceptive response is detected and rated as a nociceptive response. On the other hand, there might be nociceptive responses which are misinterpreted as defence reactions due to the restraint. As the test is terminated when a response is shown, different cut-off times do pos-

sibly not make a difference in severity with known mouse strains, since the animal is not exposed to the stimulus a longer period as necessary. Obviously, this cannot be applied to strains which cannot show nociceptive responses due to motor functional abnormalities, strains with a known lower sensitivity to nociceptive stimuli or an unknown sensitivity, as tissue damage would possibly be the outcome. It is debatable, if shorter testing intervals are more severe, since the animal has gotten a shorter time to recover from stress. In case of the unilateral hot plate test, other methods of restraint could be investigated to reduce stress and error rates. In case of the classical hot plate test, fore paw licking as a nociceptive response needs still to be investigated to reduce error rates in nociceptive response rating by the experimenter. It must be said that no investigations about error rates in the evaluation of nociceptive behaviour by the experimenter during the hot plate test have been found.

5.2.6.4. Plantar (Hargreaves) Test

This test is used to determine nociception. It is similar to the tail flick test described in 4.1.9.2, but instead of the mouse's tail, its hindpaw is used (Carter and Shieh 2015, Hargreaves et al. 1988). As the animals expierience further stress due to the fixation, we suggest categorising the test as either "mild" or "moderate".

5.2.6.5. Randall-Selitto Test/Paw-Pressure Test

This test is used to evaluate pain responses. The apparatus, which is used, is a Randall-Selitto electronic algesimeter. The mouse is immobilised and a hind paw or the tail is placed into the apparatus, to apply an increasing mechanical force, until an escape reaction is observed. Withdrawal latency is measured and evaluated (Santos-Nogueira et al. 2012).

In most of the protocols the same apparatus, a Randall-Selitto analgesiometer produced by different companies, is used. The reviewed papers describe testing on the dorsal surface of the hind paw (Fajrin et al. 2017, Lopes et al. 2017, Minett et al. 2014), the dorsal surface of the paw not mentioning if fore or hind paw (Zulazmi et al. 2015), the dorsal or plantar surface of the fore paw or hind paw (Santos-Nogueira et al. 2012), the base of the tail (Drel et al. 2007, Vareniuk et al. 2008) or the tail (Minett et al. 2014, Zhang et al. 2019).The maximal applied pressure variates between 120 g (Lopes et al. 2017), 200 g (Zulazmi et al. 2015), 250 g (Drel et al. 2007, Fajrin et al. 2017, Minett et al. 2014, Santos-Nogueira et al. 2012, Vareniuk et al. 2008, Zhang et al. 2019) or 500 g (Minett et al. 2014) with a linearly increasing pressure rate of 10 g (Drel et al. 2007). Restraint by the experimenter was described in two

protocols (Lopes et al. 2017, Santos-Nogueira et al. 2012), restraint by plastic tube in one protocol (Zhang et al. 2019). The trials are repeated three times (Drel et al. 2007, Vareniuk et al. 2008, Zulazmi et al. 2015) or four times (Zhang et al. 2019) with an inter trial interval of 5-10 min (Zhang et al. 2019), 15 min (Drel et al. 2007, Vareniuk et al. 2008) or seven days (Zulazmi et al. 2015). Last, one protocol also describes 12 h of fastening overnight (Fajrin et al. 2017).

In the Randall-Selitto test animals get restrained either by a restrainer or by the experimenter's hand. Further, pressure is applied until the animal show an escape reaction, which could be very subjective depending on the experimenter. It should be taken into consideration that the experimenter's latency time until noticing the mouse's response, processing the information and turning off the apparatus can be subject to variabilities. Therefore, the test might be categorised as "mild". Depending on the repetition of trials, the length of the inter trial time and also the maximum g used, the impact on the tissue tested needs to be investigated. This means, a higher maximum force used, and a higher number of repetitions eventually might lead to tissue damage, as the same spot is stressed several times and therefore lead to a higher severity.

5.2.6.6. Tail Flick Test

The mouse is restrained, and its tail positioned onto a photo stimulus detector. Latency time is measured until the animal retracts its tail as a response to the applied heating ray. The photo stimulus detector is switched off after a determined cut off time (Deuis et al. 2017).

To perform the tail flick test, the mouse must be restrained. This is done with the help of a plastic tube or by the experimenter's hand (Bannon and Malmberg 2007, Deuis et al. 2017, Eide et al. 1988, Nakamoto et al. 2010). Depending on the protocol, the photo stimulus is pointed on the mouse's tail with a distance of 0.4-2.4 cm (Wen et al. 2009), 1.0-2.0 cm (Nikoui et al. 2016), 1.5 cm (Bannon and Malmberg 2007, Eide et al. 1988), 3.0 cm (Dogrul et al. 2007) or 5.0-8.0 cm from the tip (Keyhanfar et al. 2013) or 3.0 cm (Langford et al. 2010) and 3.5 cm (Lichtman et al. 1993) from the base. The beam intensity of the used apparatus is set at 4.0 (Wen et al. 2009) or 5 % of the maximum output of the apparatus (Langford et al. 2010) but it is adjusted to a baseline latency time of retraction from 2-4 s (Keyhanfar et al. 2013), 2-2.5 s (Lichtman et al. 1993), 2-3 s (Dogrul et al. 2007), 2.5-3 s (Nakamoto et al. 2010), 3-4 s (Bannon and Malmberg 2007) or 4 s (Langford et al. 2010). Different cut-off

times with a maximum of 6 s (Dogrul et al. 2007), 10 s (Bannon and Malmberg 2007, Eide et al. 1988, Keyhanfar et al. 2013, Lichtman et al. 1993, Nakamoto et al. 2010, Wen et al. 2009), 12 s (Nikoui et al. 2016) or 13 s (Lichtman et al. 1993) are described as well. Finally, the measurement is repeated several times in different time intervals depending on the drug which is tested.

The tail flick test causes a nociceptive response through radiant heat, which is inescapable for the mouse. Furthermore, the animal must be restrained, a procedure which is also known to cause stress in mice (Clarkson et al. 2018, Hurst and West 2010, Pawlyk et al. 2008, Wilson and Mogil 2001). Because restraint by hand and not only handling is mentioned, it might be possible that stress level and anxiety are higher. Further, particular investigations to assess the difference between restraint by hand and restraint with a tube are necessary. Another point is the different location respective the mouse's tail, which is stimulated by the radiant heat light. It is shown, that there is no significant difference in latency time in different locations, but that there is a difference between the pigmentation of the tail (Wen et al. 2009). Unpigmented tails have a shorter latency time, which suggests that unpigmented tails are more sensitive. As the exactly used intensities of the photo stimuli are not mentioned in various protocols, a comparison is difficult. It is described that different baseline latencies are adjusted, suggesting that lower latencies work with a more intense stimulus. Most of the latencies are observed between 2-4 s and just in one case between 5-6 s. All these findings support the view that the test should be considered as either "mild" or "moderate". Another point is the different cut off times, which vary between 6-13 s. For their assessment, they should be evaluated together with baseline latency times, since it is possible that a short cut-off time with a short latency time might be less severe as a long cut-off time with a short latency time. This is because the impact of the radiant heat is longer in a long cut-off time if the animal does not show retraction within the cut-off time. It should be taken into consideration, that more repetitions of the test could mean a higher impact on the tail's tissue, for the reason that the photo stimulus is applied more often.

5.2.6.7. Tail Immersion/Hot-Water Tail Flick/Cold-Water Tail Flick Test

This test is used to determine the thermal response threshold of the animals. For this purpose, the animals are immobilized with a Plexiglas tube or by wrapping them in a diaper (Ramabadran et al. 1989), then the tail is immerged in a container with either heated water or ice water. Water temperatures are described with 4 °C, 45 °C, 50 °C, 55 °C (Luttinger 1985)

or from 50 °C heated up to 55 °C (Ramabadran et al. 1989). The time until retraction of the tail is measured and defined as thermal reaction latency (Crawley 2000, Luttinger 1985, Saha et al. 2013). As animals experience stress due to the fixation, we suggest categorising the test as “mild”.

5.2.6.8. Writhing Test

The test is used to determine abdominal, peripheral nociception and therefore the evaluation of analgesic effects of drugs. The animals are held and fixated in the neck, then injected intraperitoneally with irritants. This induces the writhing response, which is characterised by contractions of the abdominal musculature followed by a stretching of the hind legs. The number of bends is counted during the observation period (Gawade 2012). As animals are restrained and treated, they experience pain and fear, so we suggest categorising the test as either “mild”, “moderate” or “severe” dependend on the used drugs and repetitions.

5.3. Learning and Memory

The abilities of learning and memory can be evaluated with the help of different tests, which respond to different aspects of learning and memory. This includes spatial learning, associative learning, recognition memory, operant learning and motor learning. The different aspects with different tests are described in the following.

5.3.1. Spatial Learning

5.3.1.1. Barnes Maze

In this test spatial learning and memory formation is assessed. It consists of a circular, elevated platform with a number of several holes around the perimeter. Under one of the holes, an escape tunnel is attached, through which the animal can escape the bright area into a dark shelter. During several training trials the mouse learns where the escape tunnel is located. Geometrical visual cues around the platform help to locate the hole and memorise its location (Barnes 1979, Harrison et al. 2006, Pitts 2018)

The Barnes maze consists of a circular platform sized differently depending on the protocol. Its diameter measures 69 cm (O'Leary and Brown 2009, 2012), 90 cm (Harrison et al. 2006), 92 cm (Bernardo et al. 2009, Ghafari et al. 2011, Patil et al. 2009) or 122 cm (O'Leary and Brown 2009). One of the protocols has added a small white plastic wall of 15 cm height to the platform (O'Leary and Brown 2009). Further, the holes on the platform count twelve (Ber-

nardo et al. 2009, Harrison et al. 2006), 16 (O'Leary and Brown 2009, 2012) or 20 (Ghafari et al. 2011, Patil et al. 2009). The diameter of the holes was 5 cm in every protocol except two, where it measures 4.45 cm (O'Leary and Brown 2009, 2012). The Barnes maze is elevated above the floor 48.4 cm (O'Leary and Brown 2009, 2012) or 56 cm (Harrison et al. 2006). The target box under one of the holes measures 28 x 22 x 21 cm (Ghafari et al. 2011, Patil et al. 2009), 8 x 8 x 8 cm (Harrison et al. 2006) or 13 x 29 x 14 cm (O'Leary and Brown 2009, 2012) and is connected with the particular hole through a transparent plastic tube, which is 50 cm long and 5 cm in diameter (Ghafari et al. 2011, Patil et al. 2009), a white escape ramp (Bernardo et al. 2009), a white acrylic ramp (Harrison et al. 2006), a black metal ramp covered with 1.4 cm wire mesh (O'Leary and Brown 2009) or a wooden, black painted staircase measuring 14 x 10 cm with 5 steps á 1.5 cm length 10 cm width and 2 cm height, with the top staircase painted white (O'Leary and Brown 2012). The animals are allowed to stay in the escape box for 2 min during the pretraining phase and 1 min during a trial (Ghafari et al. 2011, Patil et al. 2009) or 30 s (Harrison et al. 2006, O'Leary and Brown 2009, 2012). Trials are proceeded as the following: four trials per day in four consecutive days (Ghafari et al. 2011, Patil et al. 2009), training sessions of four trials on five consecutive days (Bernardo et al. 2009, Harrison et al. 2006), followed by one probe trial (Harrison et al. 2006), habituation trial in squads of 3-4 mice, conducted underneath an inverted transparent 2 L glass breaker placed directly next to the escape hole, followed by an acquisition training phase of four days with four trials per day and finally one probe trial without escape box and without buzzer followed by two reversal training trials and another probe trial without visual cues (O'Leary and Brown 2009) or four habituation trials conducted the same way as described before (O'Leary and Brown 2009) followed by an acquisition trainings phase of 15 days with two trials per day, followed by an acquisition probe trial phase of four days, then two probe trials without escape box and four probe trials with escape box, two reversal training trials, one probe trial without visual cues and another two reversal training trials (O'Leary and Brown 2012). Trial time is described with 3 min (Ghafari et al. 2011, Patil et al. 2009) or 5 min (Harrison et al. 2006, O'Leary and Brown 2009, 2012). Inter trial time is further conducted with 10-15 min (O'Leary and Brown 2009, 2012), 15 min (Ghafari et al. 2011, Patil et al. 2009) or 15-20 min (O'Leary and Brown 2009). Some of the protocols use a start box described as a black start box of 13 x 13 x 13 cm (Bernardo et al. 2009, Harrison et al. 2006) which was lifted after 30 s (Harrison et al. 2006), a blue polyvinyl chloride start tube with 12.5 cm height and 8 cm diameter (O'Leary and Brown 2009, 2012) which is lifted after 10 s (O'Leary and Brown 2012) or 5-10 s (O'Leary and Brown 2009) or a black coloured cylindrical start chamber of

10.5 cm which is lifted after 10 s (Ghafari et al. 2011, Patil et al. 2009). Most of the protocols use aversive stimuli like a buzzer with 85 dB which is turned off immediately after mouse enters the escape box (Ghafari et al. 2011, Patil et al. 2009), bright lights (Harrison et al. 2006), two 150 W flood lamps 155 cm above the maze and a buzzer with 1-37.2 kHz and 89 dB placed 15 cm above maze (O'Leary and Brown 2012) or a buzzer with 0-37.2 kHz and 86 dB placed 20 cm above the maze and two lights of 150 W placed 76 cm above the maze (O'Leary and Brown 2009). The buzzer is turned on in acquisition training trials until the mouse escapes (O'Leary and Brown 2009, 2012). Furthermore, protocols describe guidance to the escape box during training trials (Ghafari et al. 2011, Patil et al. 2009), or that if the animal fails to escape into the box it is either guided with a 500 ml plastic tube or placed directly into the box (O'Leary and Brown 2012) or coaxed by the experimenter through touching the mouse's tail (O'Leary and Brown 2009). Visual cues around the maze or room are described in all protocols.

The Barnes Maze takes place in an open spaced environment which leads to anxiety behaviour in mice and causes stress (Bourin et al. 2007, Prut and Belzung 2003, Wilson and Mogil 2001). Additionally, aversive stimuli like bright light and a buzzer noise are used, which possibly lead to a higher stress level in mice. A positive factor is, that the animal can escape these aversive stimuli except for probe trials. Since data suggests that these animals fear heights (Stowers et al. 2017), the elevation of the platform might have influence on the animal's stress level. These findings support the suggestion to categorise the test as at most "mild". Furthermore, guiding the animal to the escape hole by touching its tail, leading it with the hand or by putting it into the box are likely to be more stressful, as guiding the mice with the help of a plastic tube. This is plausible because data shows that handling with transparent plastic tubes causes lower stress and anxiety as holding the mouse with the hand by the tail (Gouveia and Hurst 2013, 2017, Hurst and West 2010). Trial time might have an impact on the mice's welfare, since longer trial times could potentially lead to familiarisation and habituation and therefore decrease the animal's stress level. However, it is unclear what trial length and how many repetitions would be required to reduce stress and anxiety significantly. Studies indicate that repeated open field testing has no negative impact on the animal's welfare (Bodden et al. 2018), which should be valid for the Barnes Maze as well.

5.3.1.2. Cheese Board Test

The test evaluates spatial learning ability and memory. The apparatus consists of an elevated platform with eight rows of four holes radially collocated around the platform. A visual cue in form of a flag was located at the rewarded hole and helped to locate and memorise the hole. To reduce the olfactory influence, a film of the rewards smell is applied to each hole. Several training sessions, including pretraining introducing the cheeseboard, trials with a marked hole and trials without the flag at the hole are performed. Finally, the test trial is done without reward or visual cue and the time, the mouse spends in the zone where the reward was located, is measured. Another trial is done with rotating the cheeseboard 180°. If the animal cannot find the reward within 2 min through training sessions, it is placed or guided next to the hole (Llano Lopez et al. 2010). As this procedure is performed voluntarily by the animal, the test should be regarded as with "no burden".

5.3.1.3. Delayed Matching-to-Position and Delayed Non-Matching-to-Position Performance (DMTP/DNMTP)

DMTP/DNMTP is used to evaluate spatial working memory. An operant chamber is used for testing, that uses a stimulus which has to be held in mind. This can be, for example, nose-poke response holes paired with a cue lamp (Alexandra Bernardo et al. 2007, Yhnell et al. 2016) or retractable levers (Goto et al. 2010). Normally, a reinforcement stimulus is used as well. First, the animal has to make a response to one of the spatial locations. In the following, it must choose between the other locations depending on the first choice. Depending on the task, "matching" or "non-matching", the mouse has to choose the same or the opposite location in the following trials. Finally, different delay times are introduced between the initial response and the presentation of the choice phase (Alexandra Bernardo et al. 2007, Goto et al. 2010, Yhnell et al. 2016). As this procedure is performed voluntarily by the animal, the test should be regarded as with "no burden".

5.3.1.4. Morris Water Maze

This task is used to investigate spatial learning ability and memory. The mouse swims in a water-filled tank, trying to find a platform, installed and hidden just below the water surface. Geometrical visual cues around the tank help to locate the platform and memorise its location. The tank is circular, varying in diameter and water depths, depending on the laboratory protocol used. The water is warmed to 25 °C of temperature. Several training sessions, including pretraining introducing the maze, visible platform and hidden platform trials are re-

peated so the animal can memorise the location of the platform. At the end, the experiment is done without a platform. The time the animal spends in the zone in which the platform was located previously is measured. The task can also be used to evaluate reversal learning by placing the platform in a new location.

Different protocols of the Morris water maze were reviewed. A difference is the used diameter of the tank, which is 84 cm (Shelton et al. 2008), 92 cm (Bernardo et al. 2007) 105 cm (Lijam et al. 1997), 107 cm (Bernardo et al. 2009), 120 cm (Costa et al. 2001, Higaki et al. 2018, Janus 2004, Sakata et al. 2009), 122 cm (Patil et al. 2009), 140 cm (Qu et al. 2009), 150 cm (Bromley-Brits et al. 2011) or 154 cm (Mehla et al. 2019). Height of the wall varies between 45 cm (Qu et al. 2009), 47 cm (Janus 2004), 48 cm (Mehla et al. 2019) and 76 cm (Patil et al. 2009). Water temperature is described with 21±1 °C (Patil et al. 2009), 22 °C (Bromley-Brits et al. 2011), 22±1 °C (Mehla et al. 2019), 23±2 °C (Higaki et al. 2018, Sakata et al. 2009), 24 °C (Shelton et al. 2008), 24-25 °C (Janus 2004) 27 °C (Costa et al. 2001) or 30 °C (Qu et al. 2009). The water colour is either opaque by the use of non-toxic white paint (Bromley-Brits et al. 2011, Costa et al. 2001, Janus 2004, Mehla et al. 2019, Shelton et al. 2008) or transparent. The used platforms during training trials have a diameter between 10 cm (Bromley-Brits et al. 2011, Janus 2004), 11 cm (Costa et al. 2001), 12 cm (Shelton et al. 2008), 15 cm (Qu et al. 2009) or 10 cm² (Bernardo et al. 2007, Bernardo et al. 2009), 6 x 6 cm (Higaki et al. 2018, Sakata et al. 2009) as well as 12 cm radius (Mehla et al. 2019). The platform is submerged below the water's surface 0.5 cm (Janus 2004), 0.5-1 cm (Mehla et al. 2019), 1 cm (Bernardo et al. 2007, Bernardo et al. 2009, Costa et al. 2001), 1.5 cm (Higaki et al. 2018, Patil et al. 2009, Qu et al. 2009, Sakata et al. 2009, Shelton et al. 2008) or placed 1 cm above the surface (Bromley-Brits et al. 2011). The platform is either visible through a clue placed on the platform (Bernardo et al. 2007, Bernardo et al. 2009, Bromley-Brits et al. 2011, Costa et al. 2001, Janus 2004, Shelton et al. 2008) or invisible (Bernardo et al. 2007, Bernardo et al. 2009, Bromley-Brits et al. 2011, Higaki et al. 2018, Lijam et al. 1997, Mehla et al. 2019, Patil et al. 2009, Qu et al. 2009, Sakata et al. 2009). In all protocols, the probe trial is conducted without a platform. Trial time varies between 60 s (Bernardo et al. 2007, Bernardo et al. 2009, Bromley-Brits et al. 2011, Costa et al. 2001, Janus 2004, Lijam et al. 1997, Mehla et al. 2019, Patil et al. 2009, Shelton et al. 2008), 90 s (Qu et al. 2009) or 120 s (Higaki et al. 2018, Sakata et al. 2009). If the animal reaches the platform it is allowed to stay there 5 s (Bromley-Brits et al. 2011), 10 s (Higaki et al. 2018, Sakata et al. 2009), 15 s (Qu et al. 2009) or 20 s (Janus 2004). If the mouse doesn't find the platform it is put there

(Higaki et al. 2018, Janus 2004, Mehla et al. 2019, Sakata et al. 2009) and allowed to remain 15 s (Shelton et al. 2008), 20 s (Bromley-Brits et al. 2011) or 30 s (Patil et al. 2009). The different procedures are described with visible flag trials of four trials in one day (Shelton et al. 2008), four trials per day for five days (Bernardo et al. 2007) or five trials in one day (Bromley-Brits et al. 2011). Then hidden platform trials are conducted the following: four trials the same day as visible platform, eight trials the next day, then five trials the next day and three trials after surgery (Shelton et al. 2008), one trial per day for five days (Qu et al. 2009), two trials per day for 14 days (Costa et al. 2001), twelve trials per day in blocks of four trials for four consecutive days and two probe trials without platform (Lijam et al. 1997), five trials per day for five days (Higaki et al. 2018, Sakata et al. 2009), four trials per day for six days and two probe trials in one day (Janus 2004), five trials per day for four days, then one probe trial on one day without platform (Bromley-Brits et al. 2011), four trials per day for eight days then one trial without platform and afterwards five blocks á four trials in one day (Mehla et al. 2019), four trials per day in four days (Patil et al. 2009) or four trials per day for nine days and afterwards four trials on one day without pause and a probe trial without platform (Bernardo et al. 2009). Inter trial time is described with 5 min (Shelton et al. 2008) 20 min (Higaki et al. 2018, Mehla et al. 2019, Sakata et al. 2009), 20-30 min (Janus 2004), 60 min (Costa et al. 2001) or no inter trial time (Bernardo et al. 2009). Also, an inter block interval of 15-20 min (Mehla et al. 2019) is described. In one protocol a procedure is used in which the animal is placed onto the platform first, and then step by step departed from the platform and put in the water (Janus 2004).

Mice tend to escape from a water environment as fast as possible (Crawley 2000), which suggest that they are possibly not comfortable with swimming or staying in the water. Even if mice can escape the water in training trials by locating and entering the platform, it has to be taken into consideration, that the platform is removed during some trials. It seems possible that a visible, directly marked platform might reduce stress, since animals could be able to detect the platform in less time and therefore decrease their swimming time. Size of the water tank and platform might probably influence the animal’s welfare as well. This is because a greater diameter, as well as a smaller platform, could both lead to higher swimming distances. Clearly, these parameters have to be seen in correlation, since a greater diameter paired with a bigger platform could lead to equal swimming distances as a smaller diameter paired with a smaller platform. It is possible that the length of a trial, the number of repetitions and the inter trial time may influence the animal’s stress level and welfare because of exhaustion

and recovery time. Thus, a longer trial time, a higher number of repetitions and a shorter inter trial time should possibly be categorised with increased severity. All these findings support the view that the test can be categorised as either “mild”, “moderate” or “severe”. As already mentioned before regarding the apparatus, the named parameters have to be evaluated in combination.

5.3.1.5. T-Maze/Y-Maze Delayed Alternation

This test is used to assess spatial working memory. The animal learns to find a reward on the right or left arm of a T/Y-shaped maze on several training trials. Then in the first test trial, the mouse is placed at the end of the start arm and must choose one of the arms to get a reward. After making its choice, the subject is removed and after a variable delay in the range of 30 s to 5 min, is returned in the start arm. In this second trial, the reward is now located in the opposite arm to which chosen during the first trial. The mouse has to remember which arm was baited before, to make a different choice than its first one to get the reward. As the animal is motivated by food, a feed withdrawal is done during test preparation (Deacon and Rawlins 2006). Therefore, we suggest categorising this test as either with “no burden” or “mild” depending on the duration of the feed withdrawal.

5.3.2. Associative Learning

5.3.2.1. Active Avoidance

The apparatus used, consists of two chambers divided by a wall and an electrifiable mesh floor. One chamber is dark and the other one made of transparent Plexiglas, the electrifiable floor is disconnected on the transparent side. The mouse learns to escape from the preferred dark chamber to avoid an aversive electrical stimulus. It is placed in the dark chamber of the apparatus, then a foot shock is released until the animal enters the light part. 24 h later the mouse is placed in the dark chamber and latency time until the animal enters the transparent part is measured (Bovet et al. 1969, Crawley 1999, Crawley 2000). Since the used foot shock causes pain and fear in the animals, this test should be categorised as either “mild” or “moderate” according to the current intensity used.

5.3.2.2. Conditioned Fear

This task is used to evaluate the ability of learning and remembering the association between an environmental stimulus and an aversive experience (Crawley 2000, Harro 2018). The apparatus consists of a compartment with an electrifiable floor, a shock generator and a sound

source. In the first phase, the animal is placed in the chamber and exposed to foot shocks combined with a timely overlapping auditory stimulus. The measured time showing anxiety behaviour is defined as unconditioned fear. Approximately 24 h later, the contextually conditioned fear is measured during a second run in the same chamber but without any aversive stimuli. In the following, an altered context is created by using a different chamber, odour, floor, etc. During this trial, contextual discrimination of fear conditioning is quantified by comparing anxiety behaviour with previously shown behaviour. Last, the aversive stimulus is applied to the altered context and anxiety behaviour measured and defined as cued conditioning (Boulton, Alan A., Baker, Glen B., Martin-Iverson, Mathew T. 1991, Harro 2018). Some protocols show variations of the given description.

The setup of the fear conditioning consists of different context chambers, most commonly two per protocol as one is used for the conditioning training and one as an alternated context for assessing discrimination (Chaudhury and Colwell 2002, Guzmán et al. 2009, Milanovic et al. 1998, Ponnusamy et al. 2005, Radulovic et al. 1998). However, some protocols also describe the use of three different chambers (Arp et al. 2016, Kamprath and Wotjak 2004). The chambers measure 35 x 20 x 20 cm (Guzmán et al. 2009, Milanovic et al. 1998, Radulovic et al. 1998), 20.3 x 15.9 x 21.3 cm and 30.5 x 24.1 x 21 cm (Ponnusamy et al. 2005), 28 x 21 x 22 cm and 26 x 31 x 21 cm (Chaudhury and Colwell 2002), 30 x 24 x 26 cm or a round chamber with a diameter of 30 cm (Arp et al. 2016), 26 x 21 x 10 cm (Logue et al. 1997). Furthermore, a chamber with cubic shape, a cylinder and a chamber with hexagonal prism shape are described (Kamprath and Wotjak 2004). In some protocols, the chambers are surrounded by a particular context box that could measure 58 x 30 x 27 cm (Milanovic et al. 1998, Radulovic et al. 1998) or an igloo ice chest of 54 x 30 x 27 cm (Logue et al. 1997). In most of the protocols, the chamber is made of transparent Plexiglas (Guzmán et al. 2009, Logue et al. 1997, Milanovic et al. 1998, Radulovic et al. 1998), one it is a shuttle-box compartment (#ENV-010MC; Med Associates) with transparent front and back walls (Ponnusamy et al. 2005), a chamber with aluminium sidewalls and Plexiglas rear wall, ceiling and hinged front door (Chaudhury and Colwell 2002) or made of two metal walls and two Plexiglas walls or opaque side walls with a rough surface and a back plane made of Plexiglas (Kamprath and Wotjak 2004). For creating different contexts, protocols additionally mention the chamber to be bisected by a diagonal clear (Logue et al. 1997) or opaque Plexiglas partition (Guzmán et al. 2009, Milanovic et al. 1998, Radulovic et al. 1998), the usage of a 12 V light and a 20 V light (Milanovic et al. 1998, Radulovic et al. 1998) or 0.3 lx house light and 12 lx stimulus light

(Kamprath and Wotjak 2004) or markings on the wall but also use of a total different room and experimenter, where the chamber is located (Arp et al. 2016). The floor consists of a grid floor (Arp et al. 2016, Kamprath and Wotjak 2004) made of metal rods with a 1.5 mm diameter spaced apart 1.05 cm (Logue et al. 1997), with a diameter of 4 mm spaced apart 0.9 cm (Guzmán et al. 2009, Milanovic et al. 1998, Radulovic et al. 1998), with a diameter of 4 mm spaced apart 0.4 cm (Chaudhury and Colwell 2002), with a diameter of 3.2 mm spaced apart 0.8 cm (Ponnusamy et al. 2005) or a metal grid (Marsch et al. 2007). All of these floors are connected to a shock generator and serve for the fear conditioning. To alter the context, floors were varied like with white Plexiglas inserts covering the grid (Ponnusamy et al. 2005), no rods used on the floor (Milanovic et al. 1998, Radulovic et al. 1998), a flat piece of plastic (Logue et al. 1997), a flat floor (Arp et al. 2016), sawdust layer on the floor (Arp et al. 2016, Kamprath and Wotjak 2004) or silica sand used as bedding (Kamprath and Wotjak 2004).
Finally, to force the differences between the chambers, different odours are used as well. This could be, by using different cleaning solutions like 10 % ethanol (Ponnusamy et al. 2005), 70 % ethanol (Arp et al. 2016, Kamprath and Wotjak 2004, Milanovic et al. 1998, Radulovic et al. 1998), 80 % ethanol (Logue et al. 1997), 1 % acetic acid (Arp et al. 2016, Kamprath and Wotjak 2004, Milanovic et al. 1998, Radulovic et al. 1998), 10 % methanol (Ponnusamy et al. 2005), water containing isoamyl acetate (Kamprath and Wotjak 2004) or putting a few drops of orange extract into a disposable cup (Logue et al. 1997). The procedure is described with the first phase of fear conditioning. Some protocols mention a habituation time of 2 min (Ponnusamy et al. 2005), 3 min (Arp et al. 2016, Chaudhury and Colwell 2002, Kamprath and Wotjak 2004, Marsch et al. 2007) or 10 min (Logue et al. 1997) prior to the conditioning. Afterwards, a context exposure of 3 min (Guzmán et al. 2009, Milanovic et al. 1998, Radulovic et al. 1998), three phases of 2 min (Ponnusamy et al. 2005), two phases of 30 s or six phases of 30 s (Chaudhury and Colwell 2002), 20 s (Kamprath and Wotjak 2004, Marsch et al. 2007), 30 s (Arp et al. 2016, Radulovic et al. 1998), two times for 20 s (Kamprath and Wotjak 2004) or 30 s after another exploring time of 2 min (Logue et al. 1997). The exposure to the context is followed by a 2 s foot shock in all protocols, except one of only 1 s (Kamprath and Wotjak 2004) and one where the 2 s foot shock coterminates with the exposure time (Arp et al. 2016, Marsch et al. 2007). The shock has a constant current of 0.2 mA (Chaudhury and Colwell 2002, Milanovic et al. 1998, Ponnusamy et al. 2005), 0.35 mA (Logue et al. 1997), 0.4 mA (Arp et al. 2016), 0.7 mA (Guzmán et al. 2009, Marsch et al. 2007, Milanovic et al. 1998, Radulovic et al. 1998), 1 mA (Chaudhury and Colwell 2002), 1.3 mA or 2.0 mA (Milanovic et al. 1998). The context consists either of the chamber

itself combined with the previously described features or of different tones like an 80 dB white noise (Ponnusamy et al. 2005), an 80 dB 2.8 kHz tone with a 70 dB background noise during acquisition (Chaudhury and Colwell 2002), an 80 dB, 9 kHz sine wave tone (Marsch et al. 2007), a 100 dB 2.8 kHz tone (Arp et al. 2016), a pulsed auditory signal with five counts per second with 75 dB 10 kHz (Radulovic et al. 1998), an 80 dB, 9 kHz sine wave during 10 ms raising and falling (Kamprath and Wotjak 2004) or an 80 dB train of clicks with six clicks per second and a white noise of 68 dB (Logue et al. 1997). Approximately 24 h later the contextual fear is tested by placing mice in the first context chamber for 3 min (Marsch et al. 2007, Milanovic et al. 1998, Radulovic et al. 1998), 5 min (Logue et al. 1997), or 3 min of testing every 6 h for four consecutive days (Chaudhury and Colwell 2002). For long term memory another exposure 28 days later is described (Marsch et al. 2007), as well as short-term memory testing after 30 min of conditioning (Guzmán et al. 2009). At last, testing in the alternated context is described as 3 min without the auditory stimulus and then testing for 3 min after presenting the stimulus (Logue et al. 1997, Radulovic et al. 1998) or 2 min exploring time followed by the auditory stimulus and further 6 min of animal observation. This procedure was repeated every 6 h for four consecutive days (Chaudhury and Colwell 2002). One protocol mentions the exposure to the second context approximately 24 h after conditioning with 3 min of free exploring time followed by six exposures of the auditory stimulus for 30 s (Arp et al. 2016). After another 24 h memory testing in context two or three with 3 min exploring time and 30 s of exposure to the auditory stimulus is described (Arp et al. 2016). Another protocol describes the extinction of fear by exposing the animal daily to the context for 3 min without foot shock. The day mice reached extinction criteria a foot shock reminder was carried out in a novel environment 24 h afterwards (Guzmán et al. 2009). Intertrial intervals during conditioning are described with 64 s (Chaudhury and Colwell 2002), intertone intervals with 110 s and 140 s (Kamprath and Wotjak 2004) and interstimulus intervals with 120 s (Logue et al. 1997).

As this test is used for classical and/or contextual conditioning of fear, an aversive stimulus needs to be used, which is a foot shock in this paradigm. Data suggest that neuroendocrine mechanisms change, which includes the increase in catecholamines, corticosterone, adrenocorticotropin and plasma protein (Fink 2007). Since these hormones are known as stress hormones and hence the foot shock is not escapable for the animals, we suggest categorising this test as “moderate”. Likely, longer trial times and/or higher repetition rates of this test correlate positively with the plasma corticosterone levels in mice, since studies already show

these findings in rats (Fink 2007). This is why a classification of the severity “severe” might be possible for high trial times and repetitions. Referring to the impact of foot shocks on the mice’s welfare, important factors are the current density, frequency, intensity and duration of the foot shocks as well as the contact resistance. None of the reviewed protocols mentions all of these factors and hence the impact on the animal’s body depends on the variation and combination of the particular aspects, further investigation and calculation would be required. Lastly, the inter trial time could have a positive influence on the animal’s welfare, as there might be some time for recovery the longer the intertrial time is.

5.3.2.3. Conditioned Taste Aversion

This test is a learning paradigm of classical conditioning. A pleasant novel taste (e.g. saccharin) in the drinking water is paired with an aversive stimulus via intraperitoneal injection, like lithium chloride, which causes nausea. Later the avoidance of the new taste is assessed when subsequently presented (Chambers 2018, Foy and Foy 2017). We suggest categorising this test as either “mild” or ”moderate”.

5.3.2.4. Olfactory Discrimination

Olfactory discrimination has already been described in 5.2.1.

5.3.2.5. Passive Avoidance

The apparatus used, consists of two chambers divided by a wall and an electrifiable mesh floor. One chamber is dark and the other one transparent Plexiglas, the electrifiable floor is disconnected on the transparent side. The mouse learns to stay in the less preferred, bright part of the setup, to avoid exposure to the aversive electrical stimulus in the normally preferred dark area. The animal is placed in the transparent part for 10 s, then the dark chamber is opened. As soon as the mouse enters the dark part, the door between the chambers is closed and a foot shock, that causes vocalisation and flinching, is released for 1-2 s. The mouse is left ten more seconds, to associate the stimulus with the dark chamber. 24 h later the animal is placed in the apparatus with the door open. Now latency time until the mouse enters the dark chamber is measured (Crawley 1999, Crawley 2000, Jänicke and Coper 1996). This test should be categorised as either “mild” or “moderate”, depending on the used current for the foot shock.

5.3.2.6. Radial Maze

The apparatus consists of a central start box from which 8-12 arms are radiating, and each arm contains a food reward at its end. The mouse learns to find the rewards in every arm during several training trials. During the test session, the animal starts from the central box, then time is measured until all rewards are found and/or the times the mouse enters an arm where it has already collected the food. If the animals are motivated by food, a feed withdrawal is carried out during the test preparation (Crusio and Schwegler 2005, Olton and Samuelson 1976). We suggest categorising this test as “mild” depending on the duration of feed withdrawal.

5.3.2.7. Step-Down Avoidance

This test consists of a platform placed over an electrifiable mesh floor. As soon as the animal steps down touching the mesh with all 4 paws, a foot shock is released. 24 h later the latency time to step down from the platform is measured (Borba Filho et al. 2015, Crawley 2000). Depending on the current used for foot shocks, this test should be categorised at most as “mild”.

5.3.2.8. Y-Maze/T-Maze Avoidance

This avoidance test is technically similar to the active avoidance test, but instead of a two-chamber apparatus, a Y or T-shaped maze is used. One of the arms has an electrifiable mesh floor, through which foot shocks are released (Crawley 2000). We suggest categorising this test as “mild”, depending on the current used for foot shocks.

5.3.3. Recognition Memory

5.3.3.1. Social Interaction/Recognition Test

In this test, the recognition memory is evaluated. First, a subject mouse is placed into a cage and given habituation time, to create a territory. Subsequently, an unknown stimulus mouse is introduced into the territory several times. After some repetitions, the same or a novel, unknown stimulus mouse is introduced. All interactions are recorded and then evaluated (Jacobs et al. 2016, Winslow 2003).

The different protocols describe a particular individual housing phase to establish a home cage territory of 3-5 days (Winslow 2003), 4 days (Prado et al. 2006), 7-10 days (Ferguson et al. 2000) or 10 days (Bielsky et al. 2004). An individual habituation phase in the testing cage

10 min (James et al. 2015) or 15 min prior to the experimental session (Costa et al. 2003, Kogan et al. 2000) is described. For the testing phase, protocols describe different possibilities to introduce a stimulus mouse to the subject muse. One is to just introduce the stimulus animal into the home cage of the subject mouse (Bielsky et al. 2004, Costa et al. 2003, Ferguson et al. 2000, Kogan et al. 2000, Winslow 2003), another is to place the stimulus animal inside a transparent acrylic chamber containing holes, which is then put into the subject's home cage (Prado et al. 2006) or to place an anaesthetized stimulus mouse into the testing cage (James et al. 2015). The trial time varies between 1 min (Bielsky et al. 2004, Ferguson et al. 2000, James et al. 2015, Winslow 2003), 2 min (Costa et al. 2003, James et al. 2015, Kogan et al. 2000) and 5 min (Prado et al. 2006, Winslow 2003). The inter-exposure-interval, during which the stimulus mouse is removed from the home or testing cage, is specified with 10 min (Bielsky et al. 2004, Ferguson et al. 2000, James et al. 2015, Winslow 2003), 30 min (Ferguson et al. 2000, Prado et al. 2006, Winslow 2003), 60 min, 120 min, 180 min (Winslow 2003) or depends on the particular experiment (Costa et al. 2003, Kogan et al. 2000). Exposure of the same stimulus animal and inter-exposure interval are repeated once (Costa et al. 2003, Kogan et al. 2000), twice (Costa et al. 2003, Kogan et al. 2000, Prado et al. 2006, Winslow 2003), four times (Bielsky et al. 2004, Ferguson et al. 2000, James et al. 2015, Prado et al. 2006, Winslow 2003). Some of the protocols do investigate just the mice's behaviour in the previous steps (Ferguson et al. 2000, Prado et al. 2006, Winslow 2003), but most of them subsequently introduce an unknown stimulus mouse to evaluate the behaviour and therefore the memory of the subject mouse (Bielsky et al. 2004, Costa et al. 2003, Ferguson et al. 2000, Kogan et al. 2000, Prado et al. 2006, Winslow 2003) and in one protocol the known stimulus mouse as well as an unknown stimulus mouse are introduced at the same time (James et al. 2015). Testing or home cages are described as a Perspex round-bottomed bowl (James et al. 2015) or made of plastic measuring 27 x 16 x 12 cm (Costa et al. 2003, Kogan et al. 2000).

Two animals are interacting during the social interaction/recognition test which is why both of these animals have to be seen as impacted and therefore their welfare has to be evaluated individually. Since one mouse stays in its home cage while the other mouse is introduced into the home cage, the introduced mouse is used in an unknown environment, which is furthermore the habitat of another individual. Both of the mice might likely be stressed through this procedure. The introduced animal is possibly more stressed, as the environment is unknown, and it has to be moved and handled through the experimenter (Bourin et al. 2007, Gouveia

and Hurst 2013, 2017, Prut and Belzung 2003, Wilson and Mogil 2001). However, further studies are required to investigate the exact impact on the animal's welfare. If animals are not able to interact directly but stay separated through a transparent wall, it probably decreases the severity, since no fighting is possible and injuries could be avoided. Using an anaesthetized animal might be another option to avoid fighting, but general anaesthesia might go along with side effects or risk of complications. Data suggest that there might be an effect of repeated anaesthesia on the animal's wellbeing (Hohlbaum et al. 2018). Further studies are required to evaluate the effect of habituation and reduction of stress regarding the number of repetitions and duration of trials. In general, we suggest categorising this test as with "no burden" for female mice since they normally do not defend their territory with aggressive behaviour and "mild" for male mice because they are more likely to defend their territory. If fighting is induced, the severity might change. Additionally, if an anaesthetized animal is introduced as described in James et al. (2015), the severity should be categorised as either "mild" or "moderate".

5.3.3.2. Object Recognition Test

This test assesses the recognition memory. The mouse is placed in a cage with two or more unknown objects, where it is given the opportunity to explore these objects. Then one of the objects is exchanged with a novel one. The time, the animal spends exploring the novel object is measured and compared to the time spent with the known one (Leger et al. 2013). As this procedure is performed voluntarily by the animal without causing pain, the test should be classified as with "no burden".

5.3.4. Operant Learning

5.3.4.1. Five-Choice Serial Reaction Time Attentional Task

This test is a schedule-induced operant behavioural task in which the mouse has to pay attention to five spatial locations at one time. The operant chamber contains a curved wall with holes, armed with stimulus lights. During the test, one of the hole's lights is illuminated for 0.5 s. Then, the mouse can get a food reward in this hole for 5 s by breaking a photocell beam. Measured are time and accuracy of responding, as well as the location of every hole nose poke (Lustig et al. 2013, Robbins 2002). As this procedure is performed voluntarily by the animal and does not cause pain or fear, the test should be classified as with "no burden".

5.3.4.2. Schedule-Induced Operant Tasks

These tasks are used to evaluate the operant learning and memory. A total automated operant chamber is used (e.g. "Skinner box"), in which the animal learns to get a reward by pressing a lever. The standard chamber contains two levers on the front wall and another lever on the back wall, which are armed with stimulus lights. A food or water dispenser is installed on the front wall. There are different schedules used to evaluate different aspects of behaviour. The mouse needs to be trained several weeks up to months, but in favour, complete automation eliminates most external effects and leads to high precision and control of the experiment (Buselmaier et al. 1981, McLeod 2018). Variations using nose poke holes instead of levers or using a social partner as a reward are known. (Baron and Meltzer 2001, Delcasso et al. 2007, Martin and Iceberg 2015). As this procedure is performed voluntarily by the animal, the test should be classified as with "no burden".

5.3.5. Motor Learning

5.3.5.1. Conditioning of Eyeblink

The cerebellar motor learning is measured with this test. A conditioning stimulus, which consists of a visual or auditive stimulus, is paired with an unconditioned stimulus, which provokes the eye blinking. This unconditioning stimulus consists of a 100 ms foot shock or an air puff, straight onto the cornea. The animal is fixed during this test, but there are apparatuses available which just fixe the mouse's head, letting the body unrestrained placed on a ball or treadmill, which can be moved. This probably should reduce stress. The eyeblink response is measured with a sensitive camera system (Disterhoft and Weiss 2017, Rasmussen et al. 2018). Since mice experience stress due to the restraint, we suggest categorising the test as "mild".

5.3.5.2. Rotarod Test

The Rotarod test to evaluate coordination and balance has been described in 5.1.3.3. To evaluate motor learning, the improvement in performance on several consecutive days is measured by latency to fall from the Rotarod. We suggest categorising this test at most as "mild".

5.4. Feeding and Drinking

Feeding and drinking is an essential and important life event and the most basic survival behaviour. Since these behaviours are affected by circadian rhythms, aging, anxiety, neuroen-

docrine axis and other brain functions, various investigations are conducted regarding these subjects. Some of them are described in the following.

5.4.1. Analysis of Feeding Behaviour

Microstructural behaviour is evaluated with the help of recorded feeding behaviour. This could relate to investigations about the neuroendocrine axis, drugs, transgenic manipulations, diseases, etc. by evaluating meal size, rate of consumption, intermeal interval or orofacial motor components (Azzara 2004, Bercik et al. 2009, Czyzyk 2013). Laboratory mice do not experience any stress during these observations. This is why the analysis of feeding behaviour should be classified as with “no burden”.

5.4.2. Cafeteria Diet

The cafeteria diet investigates human obesity by feeding a high calorie diet, consisting of food regularly consumed by humans over several weeks (Sampey et al. 2011). Often, additional tests like calorimetry are used or the effects of drugs are tested.

The diet, which is fed, varies depending on the protocol. Described are chocolate, biscuits and peanut butter (Zeeni et al. 2015), cheese- or bacon flavoured chips, marshmallows, peanut candy, filled and wafer cookies, sausage, mortadella and soda (Gasparin et al. 2018), digestive biscuits and almond paste (Luijten et al. 2019), chocolate crackers, wafers, marshmallows, mortadella, hot dog sausages, cheese and bacon chips, Doritos® chips, peanut candy, calf’s food jelly as well as soft drinks like guarana and cola (Leffa et al. 2015, Leffa et al. 2017). The composition of the diets is shown in table 3. The mice could select and consume the foods ad libitum (Gasparin et al. 2018, Luijten et al. 2019, Zeeni et al. 2015) and some protocols allowed access to the standard rodent diet additionally (Leffa et al. 2015, Leffa et al. 2017, Luijten et al. 2019). Two protocols use a daily changing menu for the diet (Leffa et al. 2015, Leffa et al. 2017). Feeding the cafeteria diet lasts 14 weeks (Gasparin et al. 2018), 15 weeks (Zeeni et al. 2015), 35 days (Luijten et al. 2019), from 3 weeks of age until 11 weeks of age (Bailey et al. 1986), 8 weeks (Muller et al. 2018), 17 weeks (Leffa et al. 2017) or 13 weeks (Leffa et al. 2015). Daily replaced with fresh food (Gasparin et al. 2018). Body weight was measured three times a week (Luijten et al. 2019), weekly (Gasparin et al. 2018, Leffa et al. 2015, Zeeni et al. 2015) or at the beginning and the end of the diet phase (Muller et al. 2018). Food intake was measured daily (Gasparin et al. 2018, Leffa et al. 2015), three times a week (Luijten et al. 2019) or weekly (Zeeni et al. 2015). Further measurements

conducted are body composition and metabolic efficiency by magnetic resonance imaging, indirect calorimetry in a temperature-controlled metabolic chamber (Luijten et al. 2019) or running capacity test five times per week and 48 h of metabolic cages (Muller et al. 2018).

Table 3: Composition of the used diets

Reference	Particular used Food	% Kilocalories Fat	% Kilocalories Carbohydrate	% Kilocalories Proteins
Muller et al. 2018		18.8	55	14.8
Zeeni et al. 2015	Chocolate	47	47	6
	Biscuits	40	55	5
	Peanut Butter	75	9	17
Gasparin et al. 2018		17	73	10
Bailey et al. 1986		7	83	10
Leffa et al. 2015, Leffa et al. 2017		51.84	40.71	7.44

The cafeteria diet uses high calorie food, which could be considered as unhealthy, to assess the impact on body and welfare of the animals. Animals do not experience any fear or pain, which is why we suggest rating the test basically as with “no burden”. Since studies indicate that the used foods could cause advanced liver damage, renal interstitial fibrosis and heart damage (Zeeni et al. 2015), it should be taken into consideration to classify the cafeteria diet as either “mild” or “moderate”, depending on the expected organ damage. It is likely that the test does not have a major impact on the psychological stress level of the mice but neverthe-less the physiological impact cannot be dismissed. Possibly, the negative impact correlates positively with the duration of feeding the diet. This should be taken into consideration while evaluating the severity of this test.

5.4.3. Conditioned Taste Aversion

Conditioned taste aversion has already been described in 5.3.2.3.

5.4.4. Conditioned Taste Preference

First, mice are singly housed and adapted to a drinking solution for a few days. Then they undergo surgery, to place a catheter subcutaneously from the neck, through abdominal muscles, into the stomach. After a few days of recuperation time, the conditioning trials are performed. The mice are placed in infusion cages, which consist of transparent walls, a mesh floor and two drinking bottles on the front wall. The infusion system is installed above the cage and output connected to the mouse's catheter. Different flavoured liquids are offered as a positive or negative conditioning stimulus. Then drinking was observed automatically, starting the auto-controlled infusion pump with either water or maltodextrin, depending on the used conditioning stimulus (e.g. three days positive stimulus, three days negative stimulus). After this phase, a two-bottle choice test measures the amount of consumed drinking solution (Sclafani and Glendinning 2003). The conditioned taste preference should be categorised as either "moderate" or "severe" because animals first undergo a surgical procedure and then are fixated due to the catheter system.

5.4.5. Restricted Daily Access

The daily consumption of water and food is restricted to a few hours per day. Normally these restrictions are used to evaluate the effects of short-acting medications, but they can also be used in other experiments (Hatori et al. 2012). We suggest classifying this test as either "mild" or "moderate", depending on the duration of water and food deprivation as well as used medication.

5.4.6. Specialised Diets

Specialised diets are used to evaluate and study gene functions, preferences and effects of different nutrients, as well as effects of hormones or drugs on feeding behaviour and weight (Crawley 2000). These diets can be high or low in particular nutrients or might be substituted by dietary supplements (Kim et al. 2016, Nuzzo et al. 2018). Since studies show that some diets could lead to organ damage (Zeeni et al. 2015), this test should be categorised as with either "no burden", "mild" or "moderate", depending on the organ damage which is expected.

5.4.7. Twenty-Four Hour Consumption

The consumption of water and food within 24 hours is measured. Therefore, food is weighed every 24 h. To also take account of spillage, the cage needs to be scanned for food crumbs lying around. Laboratory mice do not experience any stress during observations. This is why the twenty-four-hour consumption should be classified as with "no burden".

5.5. Reproductive Behaviours

Reproductive behaviours are observed to evaluate the function of complex feedback mechanisms regulated by central and peripheral hormones. Among other things, testing includes observation of courtship and copulation, nesting, parenting, pup retrieval and pup ultrasonic vocalisation as described in the following.

5.5.1. Courtship and Copulation

The individual components of sexual behaviours during different tests are recorded and then quantitated by an observer (Bean et al. 1986, Matsumoto and Okanoya 2016, Pomerantz et al. 1983, Watt 1931). Laboratory mice do not experience any stress during observations. This is why we suggest categorising the test as with "no burden".

5.5.2. Nesting (Nest Building)

The testing and evaluation of nest building can indicate distress pain and suffering in the animals (Gaskill et al. 2013, Gjendal et al. 2019, Jirkof et al. 2013). The mice are provided with nesting material and allowed to interact with the material for 24-48 h. Afterwards, the nest is removed and evaluated, which includes weight and different quality parameters. Then the mice are provided with new nesting material. In different protocols the nesting materials can variate.

For the nest building test an individual housing of the animals is provided in some of the protocols (Deacon 2006, Greenberg et al. 2016, Kraeuter et al. 2019). In others, housing in pairs (Newman et al. 2019) or groups of three animals (Gaskill et al. 2013, Gjendal et al. 2019) are described. As this test might be used to assess distress in form of medication or experimental procedures, they are used previously to the nest building test phase (Gjendal et al. 2019, Greenberg et al. 2016). They provided nesting material with a weight of 2 g (Newman et al. 2019), 2.5 g (Kraeuter et al. 2019), 3 g (Deacon 2006), 8 g (Greenberg et al. 2016), 8-10 g (Gaskill et al. 2013) or 30-35 g (Gjendal et al. 2019). The used nesting material was

Enciro-Dri®shredded paper (Greenberg et al. 2016), crinkled paper nesting material (Gaskill et al. 2013) or cotton nesting material (such as Nestlet) (Deacon 2006, Kraeuter et al. 2019, Newman et al. 2019). Some protocols mention enrichment items in usage (Gjendal et al. 2019) whereas others do not recommend the usage of these at all (Deacon 2006). Nesting material was provided at different times of the day: At 14:00 h (Gjendal et al. 2019), 15:00 h (Greenberg et al. 2016), 13:30 h (Newman et al. 2019) or 19:00 h (Kraeuter et al. 2019). The evaluation of the nests is done 10 h (day 1: 19:00 h), 24 h, 28 h, 32 h, (day 2: 9:00 h, 13:00 h, 17:00 h) and 48 h (day 3: 9:00 h) after the experimental procedure (Greenberg et al. 2016), 7-9 h after beginning of the light phase (Gaskill et al. 2013), five days later (Newman et al. 2019), at 7:00 h the next day (Kraeuter et al. 2019) or the next morning (Deacon 2006). Most protocols do not only describe the nest building behaviour, but also the time to integrate to the nest as part of the evaluation of nesting. This describes the motivation to engage in nesting behaviour when mice are provided with fresh nesting material. Therefore, one protocol provides animals with a 2.5 cm square of nestlet after 18 h (Gjendal et al. 2019), the other provides animals with a 5.08 cm square of cotton nesting material cut into quarters within the first 3 h of lights turned on (Gaskill et al. 2013). An acclimatisation period of seven days (Greenberg et al. 2016, Kraeuter et al. 2019) prior to the test beginning is described.

Nest building is conducted in a known environment, to which the animals are habituated. Additionally, it basically does not require any invasive methods. This is why we suggest classifying the test as with "no burden". Clearly, depending on the tested drugs and their application form or other experimental procedures the severity could change. Data show that enrichment items have a positive impact on the animal's welfare, stress, anxiety behaviour, etc. (Bailoo et al. 2018), that is why protocols which do not use these items should be considered as having a higher negative impact on the animals. Studies indicate that mice prefer some nesting materials over others (van de Weerd et al. 1997). As the nest has an impact on thermoregulation, preweaning mortality of pups and therefore the animal's welfare, it is likely that the nest's structure influences these factors as well.

5.5.3. Parenting

The parental care is recorded and then scored for several parameters by an observer (Brown et al. 1996, Kuroda et al. 2007, Tachikawa et al. 2013). Parenting and pup retrieval assays are often combined. Laboratory mice do not experience any stress during observations. This is why parenting should be classified as with "no burden".

5.5.4. Pup Retrieval

To assess parental care behaviour, the latency time until the parent brings back its briefly separated new-born pups into the nest and the number of pups which are retrieved is measured (Brown et al. 1996, Kuroda et al. 2007, Tachikawa et al. 2013). We suggest categorising this test as either with “no burden” or “mild”, depending on the repetitions of separating pups from their parents.

5.5.5. Pup Ultrasonic Vocalization

This test can be used to detect communication deficits or parental behaviour. New-born pups are briefly separated from their mother and placed in a box equipped with high-frequency detecting equipment. The vocalisation is then analysed and the test can be repeated over several days (Hofer 2002, Mogi et al. 2017, Okabe et al. 2013, Takahashi et al. 2009). We suggest categorising this test as either with “no burden” or “mild”.

5.6. Social Behaviours

The behavioural repertoire exhibited by mice during social behaviour tasks is observed or recorded and then evaluated. This includes different behaviours, like huddling, juvenile play behaviour, aggression or different grooming behaviours (allogrooming, barbering and whisker trimming). Nesting (nest building), social interaction, social transmission of food preference test and the three-chamber social approach task are described in the following.

5.6.1 Nesting (Nest Building)

Nesting (Nest Building) has already been described in 5.5.2.

5.6.2. Social Interaction

Social interaction has already been described in 5.3.3.1, but other tests referring to social interaction are also used in research. Two of them are described in the following.

5.6.2.1. Social Transmission of Food Preference Test

First, a demonstrator mouse is allowed to eat food mixed with a novel flavour. After consuming the novel food, the demonstrator is allowed to interact with observer mice. In the final choice phase the observer mice are given a choice between the flavour of food eaten by the demonstrator and some other novel flavour. Normal observer mice will prefer the flavour of food eaten by the demonstrator (Wrenn 2004). Laboratory mice do not experience any pain

or fear during this test. Furthermore, the animals act voluntarily, which is why we suggest classifying this test as with “no burden”.

5.6.2.2. Three Chamber Social Approach Task

This test evaluates social behaviour and interaction in mice. The apparatus consists of a transparent box, consisting of three chambers. The subject mouse is habituated first, then a novel mouse is placed under a pencil cup in one of the chambers. Behaviour, like sniffing and entering the chamber, is observed and evaluated. In a second trial the yet known mouse is placed in one chamber and another unknown mouse is placed in the other chamber, both under a pencil cup again. Social interaction behaviour is observed, measured and compared (Faizi et al. 2012, Moy et al. 2004, Pearson et al. 2010). Laboratory mice do not experience any pain or fear during observations, which is why we suggest classifying this test as with “no burden”.

5.6.3. Aggression

Aggression is not only an important behaviour in mice, as the family organisation is defended by a dominant male mouse. It is also an interesting behaviour for various animal models of human diseases (Perepelkina et al. 2017). Therefore, some of the used behavioural tasks for assessment of aggression are described in the following.

5.6.3.1. Standard Opponent

This test is used for dominance and aggression evaluation. A male mouse is chosen as the standard opponent mouse, for its repetitive dominant or submissive behaviour. The animal is paired with the mouse to be tested, observing and evaluating its behaviour for 5 min (Brain and Poole 1974, Martinez et al. 1994). Depending on duration and intensity of the fighting, this test should be categorised as either with “no burden”, “mild” or “moderate”.

5.6.3.2. Isolation Induced Fighting

This test is a variation of the standard opponent test, described in 5.6.3.1, intensifying aggressive behaviour caused by isolated housing (Hadfield et al. 1982, Valzelli et al. 1974). Depending on duration and intensity of fighting, this test should be categorised as either with “no burden”, “mild” or “moderate”.

5.6.3.3. Resident-Intruder Test

This test is a variation of the standard opponent test, described in 5.6.3.1, where the standard opponent is the intruder in the home cage of the mouse to be tested. It can be conducted with single mice or mouse families with pups in a large home cage (Koolhaas et al. 2013, Rammal et al. 2010). Depending on duration and intensity of fighting, this test should be categorised as either “mild” or “moderate”.

5.6.3.4. Round-Robin Approach

This test defines the dominance hierarchy within a male group. Every possible combination of two mice is compared for dominant and submissive behaviour (Brain and Nowell 1970, Drickamer 2001). Depending on the duration and intensity of fighting between the group members, we suggest categorising the test as either with “no burden”, “mild” or “moderate”.

5.6.3.5. Social Dominance

This test evaluates dominance and aggression behaviour without physical contact. The apparatus consists of two start boxes at the end of a 30 cm long tube and a central neutral area, which is separated by two gates. The mice are placed each at one end of the tube and then released, observing their behaviours. As they can smell each other, one mouse will show more dominant postures, while the submissive mouse will back away (Crawley 2000, van de Weerd et al. 1992). We suggest categorising this test at most as “mild”.

5.7. Emotional Behaviours

5.7.1. Anxiety- and Fear-Related Behaviours

Behavioural tests which measure anxiety- and fear-related behaviours can be classified into conditioned and unconditioned models (Crawley 2000, Sousa et al. 2006). Conditioned models include conflict based and non-conflict based tests, like the Vogel-Conflict and Geller-Seifert test or different conditioned fear tests, like ultrasonic vocalisation and startle tests. Unconditioned models are divided into exploration-, interaction-, response- and defence-based tests. Among others, they include elevated mazes, open field test, social interaction, ultrasonic vocalisation and startle responses (Sousa et al. 2006).

5.7.1.1. Conditioned Fear

Conditioned fear has already been described in 5.3.2.2.

5.7.1.2. Elevated Platform

This test is used to assess anxiety behaviour. The mouse is placed on an uncovered, bright, open platform, which is elevated above the floor. The behaviour is recorded during a predefined time and later evaluated. A modification of this test adds grip covered steep slopes to the edges of the platform, which can be explored (Ennaceur 2012). Laboratory mice do not suffer from any pain during this test but might experience fear, which is why we suggest classifying this test at most as “mild”.

5.7.1.3. Elevated Plus-Maze

This test is used to evaluate anxiety behaviours. The apparatus consists of a plus-shaped maze, elevated above the floor, with two open and bright lit arms and two covered and dark arms. The mouse is placed in the central area, then the willingness to visit the open arms and time spent there is measured (Walf and Frye 2007). Laboratory mice do not suffer from any pain during this test but might experience fear. However, the animals can escape the bright lit arms, which is why we suggest classifying this test as with “no burden”.

5.7.1.4. Elevated Zero-Maze

This test is a modification of the elevated plus-maze. The apparatus consists of two enclosed and two open arms that form a circle, elevated from the ground. The time spent in enclosed and open arms is measured and compared (Cook et al. 2002, Shepherd et al. 1994). Laboratory mice do not suffer from any pain during this test but might experience fear. However, the animals can escape the bright lit arms, which is why we suggest classifying this test as with “no burden”.

5.7.1.5. Emergence Test

This test is used to assess anxiety. An open field area, as already described in 5.1.2.3., is used. Additionally, a small plastic compartment with an exit is placed centrally into the open field. The mouse is placed under the cylinder and tested for up to 15 min. The general locomotor activity, the latency until the mouse exits from the cylinder and the time spent inside the cylinder and exploring the cylinder are measured (Miao-Kun Sun 2007). Laboratory mice do not suffer from any pain during this test but might experience fear. However, the animals can escape the bright lit area, which is why we suggest rating this test as with “no burden”.

5.7.1.6. Fear/Defence Test Battery

This battery is used to evaluate antipredator defensive behaviours. The apparatus consists of a 6 x 1 m area with rounded ends, divided in the middle by a 4 m long portion and so creating an oval runway. The floor is made of concrete, with labelled meter-distances for measurement and the walls are made of plywood. The mouse is placed into the apparatus, first observing the animal and measuring the covered distance. Then the experimenter enters the runway, getting closer to the mouse at a speed of 1.0 m/s. Avoidance and flight distance are recorded. In a second trial, the flight speed is measured by chasing the animal a determined distance at a speed of 2.0 m/s. In another trial, reactions of anxiety and defence are recorded and evaluated. For this purpose, one plywood partition is closed to create a straight ending linear runway. The experimenter enters the runway from the open side, chasing the animal at a speed of 0.5 m/s, stopping at different distance labels for 30 s. Other protocols also describe the use of a rat, as natural predator, entering the runway (Blanchard et al. 1997, Blanchard, Robert et al. 1995). We suggest categorising this test at most as “mild”, depending on the details of each protocol.

5.7.1.7. Fear-Potentiated Startle

The acoustic startle response, which has been described in 5.2.3., is paired with a foot shock (Boulis and Davis 1989, Walker et al. 2003). Therefore, we suggest categorising this test as either “moderate” or “severe”, depending on the intensity of the current used for foot shocks.

5.7.1.8. Four Plate Test

This test is used as a model of anxiety and conditioned fear through passive avoidance. The mouse is placed in a chamber, which has an electrifiable floor, consisting of four metal plates. After a habituation time, a foot shock is applied, whenever crossing from one plate to another. During a one-minute period, the number of applied foot shocks is measured. The mouse can avoid the shocks by remaining motionless (Hascoët M. 2011). We suggest categorising this test as either “mild” or “moderate”, depending on the intensity of the current used for foot shocks.

5.7.1.9. Learned Safety/Fear Inhibition

This paradigm is used to investigate the dysregulation of emotional behaviours, by inhibiting fear. The test is similar in structure to the conditioned fear, which has been described in 5.3.2.2. It only differs in the temporal pairing of the neutral and the aversive stimulus. In con-

trast to the fear conditioning, the neutral conditioning is never presented simultaneously to the aversive stimulus, but explicitly unpaired. The animal learns to associate the neutral stimulus with the absence of the aversive stimulus, thus interpreting it as a safety signal (Pollak et al. 2010). We suggest categorising this test as either ”moderate” or “severe”, depending on the aversive stimulus.

5.7.1.10. Light-Dark Test

The light-dark test has already been described in 5.2.2.1.

5.7.1.11. Novelty Suppressed Feeding

The test is used to assess anxiety-like behaviour. The setup consists of a wide-open container, filled with normal bedding and a weighted, white paper circle, where food is placed. The mouse is placed in the novel arena and time is measured until the animal eats the familiar food. A food withdrawal is necessary to motivate the animal to consume the food (Samuels B.A. 2011). Depending on the duration of food withdrawal, this test should be categorised as “mild”.

5.7.1.12. Open Field Test

The open field test, as described in 5.1.2.3., is used to describe and quantitatively evaluate the spontaneous activity, exploratory behaviour and locomotion of rodents, but also to measure fear-related behaviours. For this purpose, freezing and defecation, as well as the time the animal stays next to the corners and walls (thigmotaxis), are measured. We suggest categorising the test at most as “mild”.

5.7.1.13. Predator Odour

This test is used to assess anxiety-like behaviour. The mouse is confronted with predator odour, placed in an experimental cage setup. Then anxiety behaviour is observed and evaluated (Laska and Sievert 2016, Otsuka 2017). The animal might experience fear, but since it is not exposed to the predator directly, we suggest categorising this test at most as “mild”.

5.7.1.14. Pup Ultrasonic Vocalisation

Pup ultrasonic vocalisation has already been described in 5.5.5.

5.7.1.15. Social Interaction

Social interaction has already been described in 5.3.3.1.

5.7.1.16. Visible Burrow System

This test assesses defensive behaviour. The apparatus consists of an open surface part connected with dark tunnels and a chamber system underneath. Food and water are located in the open space area and a light-dark circle of 12 h is used. The subterranean part is lit with red light during 24 h per day and in both parts videotaping is possible. A group of mice is housed in the apparatus and normal behaviour patterns are observed and evaluated. After some days, a predator stimulus, for example a cat, is placed on the upper surface for about 15 min. Afterwards, behaviour is recorded and assessed again for several days (Blanchard, Robert J. et al. 1995). The animal might experience fear but is not exposed to the predator directly and could also hide in the dark chamber system. This is why we suggest categorising this test at most as “mild”

5.7.1.17. Vogel Conflict Test

This test is used to assess anxiety behaviour by using anxiolytic drugs. The apparatus consists of a Plexiglas box with an electrifiable mesh floor and a drinking bottle. Licking is measured automatically and after a determined number of licks a foot shock is released. The test is performed starting with a habituation trial of 40 min without foot shocks. One week later a second 40 min trial is performed after the administration of drugs, counting the foot shocks released. Before both trials a water withdrawal of two days is made (Umezu 1999). We suggest categorising this test as either “mild” or “moderate”, depending on the intensity of the used current and the duration of water withdrawal.

5.7.2. Behaviours and Models Related to Symptoms of Psychological Diseases

5.7.2.1. Marble Burying

This test evaluates compulsive-like behaviour. The animal is placed in a cage with 20 marbles placed on top of the bedding. After 30 min the buried marbles are scored. Different factors like diseases, genetics or drugs influence the burying of the marbles (Angoa-Pérez et al. 2013). Laboratory mice do not experience any pain or fear during this test, which is why we suggest classifying this test as with “no burden”. However, severity might change due to the used drugs.

5.7.2.2. Nestlet Shredding

This test evaluates compulsive-like behaviour. The animal is placed in a cage, where a pulped cotton fibre nestlet is placed on top of the bedding. After a duration of 30 min, the nestlet remaining intact is determined, by comparing weight before and after the test trial. Different factors like diseases, genetics or drugs influence the behaviour of shredding the nestlet (Angoa-Pérez et al. 2013). Laboratory mice do not experience any pain or fear during this test, which is why we suggest classifying this test as with "no burden".

5.7.3. Models of Schizophrenia

5.7.3.1. Acoustic Prepulse Inhibition

This test is used to measure sensorimotor and information gating. The acoustic startle response, which has been described in 5.2.3, is paired with a preexposure to a weaker 90 dB tone. This inhibits the response to the following 120 dB tone, if presented within 100 ms. Abnormal sensory inhibition is noticed in deficits in information gating and can also be stimulated through the use of amphetamines. Paired with the procedure already described in 5.2.3 we suggest classifying this test as either "mild" or "moderate".

5.7.3.2. Behavioural Sensitisation

This test measures the sensitisation by repeated exposure to a stimulus. This leads to an enhanced response to the same stimulus and therefore measurable behavioural changes. For example, repeated exposure to an open field test normally decreases exploration behaviour through familiarisation. However, through repeated drug administration, hyperlocomotion is observed (Jung et al. 2013, Weidenauer et al. 2018). We suggest classifying this test as either with "no burden", "mild" or "moderate" depending on the used drugs or other protocol variations, like the repeated stimulus.

5.7.3.3. Latent Inhibition

This test assesses the inhibition of stimulus-response learning and information gating. The animal is placed in an operant chamber, which contains an electrifiable mesh floor and a loudspeaker, to present a tone. In the first step, the tone is played in a neutral context. Later, the tone is matched with a foot shock. As a next step, the mouse is offered water and after some time, the tone is presented again. The drinking behaviour is then recorded for 5 min. Another animal is not introduced to the tone in a neutral context before, but the tone is immediately paired with a foot shock. Then, following the same procedure, the drinking behav-

iour is recorded and evaluated. The preexposure of the stimuli normally leads to an inhibition of later stimulus-response learning, as the stimulus is not associated with the foot shock. Thus, the first mouse will normally have lower drinking suppression as the second mouse. Abnormal inhibition is noticed in information gating deficits and can also be stimulated through the use of amphetamines (Lubow 1973). Depending on the intensity of the used current, the test should be categorised at most as "mild".

5.7.3.4. Hyperlocomotion

Hyperlocomotion after application of psychostimulant drugs can be assessed in the open field test, described in 5.1.2.3. Depending on the used drugs, we suggest classifying the assessment of hyperlocomotion as either with "no burden" or "mild".

5.7.4. Models of Depression

5.7.4.1. Immobilisation/Restraint

This test is used to evaluate and/or create stress responses. The animal is immobilised through a restraint apparatus for a particular time, with few or no possibilities to move. Depending on the protocol, the mice might be deprived of food and water.

The immobilisation or restraint of the mouse is generally ensured by a restraint apparatus, which is described as a tube (Chu et al. 2016, Huang et al. 2015, Jang et al. 2018, Marianno et al. 2017, Padgett et al. 1998), a rodent immobilisation bag (Kedia and Chattarji 2014), jars (Khandve 2013), or a rodent restraint device (Sulakhiya et al. 2016). The apparatus are made of polypropylene (Marianno et al. 2017), transparent plastic (Chu et al. 2016, Khandve 2013), Plexiglas (Sulakhiya et al. 2016) or a simply centrifuge/conical tube (Huang et al. 2015, Jang et al. 2018, Padgett et al. 1998). Measurements within the protocols are 3 cm in diameter and 11.5 cm of length (Marianno et al. 2017), 3 cm in diameter and 10 cm of length (Chu et al. 2016) or 50 ml for the centrifuge tubes (Jang et al. 2018, Padgett et al. 1998). All of the apparatuses are described as well ventilated. The reviewed protocols describe immobilisation at different levels, so complete immobilisation without any physical movements is possible (Kedia and Chattarji 2014, Sulakhiya et al. 2016), animals can move their head and forelimbs, but are not able move or turn their body and hindlimbs (Chu et al. 2016) or forward-to-backward and side-to-side mobility are prevented by inserting gauze into the tube (Jang et al. 2018). Furthermore, one protocol describes the restrainment apparatus to be placed vertically (Jang et al. 2018). In most of the protocols the animals do not have access

to food or water (Chu et al. 2016, Kedia and Chattarji 2014, Padgett et al. 1998, Sulakhiya et al. 2016), but one protocol describes access to food and water (Khandve 2013). Restrainment durations of 1 h (Marianno et al. 2017), 2 h (Huang et al. 2015, Jang et al. 2018, Kedia and Chattarji 2014), 4 h (Khandve 2013), 6 h (Kumar et al. 2010, Sulakhiya et al. 2016), 12 h (Padgett et al. 1998) or 24 h (Chu et al. 2016) are described. Restraint phases are conducted once (Chu et al. 2016, Jang et al. 2018, Kedia and Chattarji 2014, Marianno et al. 2017, Sulakhiya et al. 2016), eight times (Padgett et al. 1998), six days per week during eight weeks (Huang et al. 2015) or for 60 days (Khandve 2013). Lastly, one protocol describes the procedure of immobilisation by taping the animal's limbs on a board by using zinc oxide tape after putting the mice on their backs. For releasing the mice, tape is moistured with acetone and then unravelled (Kumar et al. 2010).

Not only studies indicate that restraint induces stress (Pawlyk et al. 2008), but immobilisation is one of the most used methods to induce stress responses in animals for further investigations (Glavin et al. 1994, Patchev and Patchev 2006). Furthermore, most of the reviewed protocols use water and food deprivation during the immobilisation phase. Since data indicate that water deprivation alters the plasma corticosterone concentration in correlation with the duration of the deprivation (Bekkevold et al. 2013) and metabolic disturbance as well as changes in hepatic metabolites are observed (Cui et al. 2015), additional factors regarding the animal's welfare are given. This is why we suggest categorising restraint/immobilisation with either a "moderate" or "severe" severity, depending on restraint duration and repetitions. Protocols in which the mouse is able to move its head and forelimbs could possibly be less stressful, but further studies for evaluating the differences between full body immobilisation and only partly immobilisation are required. It is likely that the intensity of the stress response is correlated positively with the duration of restraint, but could underlie changes as well (Glavin et al. 1994). Furthermore, it is not clear if a habituation effect occurs after repeated restraint procedures (Glavin et al. 1994).

5.7.4.2. Forced Swim Test

This test measures the listlessness in mice. The animal is placed in a glass cylinder, filled with water to a depth that it cannot balance on its tail on the bottom nor escape over the top. Water temperatures from room temperature up to 35 °C are used. During the 4-20 min trial, the time actively swimming and trying to escape is measured and compared to the time the

mouse is just floating on the water surface (Can, Dao, Arad et al. 2012, Porsolt et al. 1977). We suggest categorising this test as either “mild”, “moderate” or “severe”.

5.7.4.3. Learned Helplessness

This test assesses the effects of controllable and uncontrollable stressors. Three mice are tested at the same time, while one is exposed to an escapable stressor, one to an unescapable stressor and one to no stressor. The stressor consists of a foot or tail shock, depending on the protocol, which can be stopped by fulfilling the right response, like wheel running or entering another compartment of the test apparatus. The first and the second chamber are connected, so every time the mouse in the first chamber receives a shock, the mouse in the second chamber receives the same shock at the same time. The first mouse can then terminate the stressor, while the second mouse cannot do anything. The whole test trial lasts 1 h. Twenty-four h later, mice performance is evaluated, for example during an active avoidance task. Depending on the subject to be investigated, the two trials are repeated after application of drugs to be tested (Anisman and Merali 2001). Due to the unascapable aversive stimulus which leads mice to give up on theirselfs this test should be categorised as “severe”.

5.7.4.4. Postweaning Social Isolation

To induce stress, the mice are housed isolated from other conspecifics during early postweaning live, which can also be done under different light-dark circles (Matsumoto et al. 2019, Pietropaolo et al. 2008, Valzelli 1973). We suggest categorising this test as either with “no burden”, ”mild” or “moderate”.

5.7.4.5. Social Deprivation/Social Isolation

Mice are housed isolated from other conspecifics to induce stress (Berry et al. 2012, Delini-Stula and Vassout 1981, Valzelli 1973). We suggest categorising this test as either with “no burden”, ”mild” or “moderate”.

5.7.4.6. Sucrose Preference Tests

This test measures anhedonia, based on the preferred consumption of a sweet solution, which has a rewarding effect compared to water. This test is a two-bottle choice test (see 5.4.2.) combined with a stressor, which induces depression-like behaviour (Liu et al. 2018). This test should be categorised at most as “mild”.

5.7.4.7. Tail Suspension Test

This test evaluates listlessness in mice. The mouse is held or suspended at the tail tip in a head downward position. The time and number of attempts to actively free themselves from this situation are measured and then compared with the time of immobility. A test trial can last 6 min and more, depending on the used protocol (Can, Dao, Terrillion et al. 2012, Steru et al. 1985). We suggest categorising this test as either “moderate” or “severe” depending on trial time and repetitions.

5.7.4.8. White Noise

In this experiment, stress is induced by high intensity background noises, applied through various days (Anthony et al. 1959). We suggest categorising this test at most as “mild”.

5.8. Reward and Abuse

5.8.1. Conditioned Place Preference/Aversion

In this test, the mouse learns to associate a neutral, contextual stimulus with a positive or negative stimulus. The apparatus consists of a box with two chambers, differing in form, visual cues or lightning, which are connected by a door. During the first phase, a preexposure session, the animal can explore the apparatus for up to 15 min. In the next step, a drug is applicated and depending on its effects, used as a positive or negative stimulus. The mouse is then immediately placed in the appropriate chamber, with the door closed. This procedure can be repeated, to reinforce the conditioning effect. Twentyfour hours later the test trial is performed. The mouse is placed into the chamber and the door is left open. The time spending in the reinforced compartment, in comparison with the other compartment, is measured. Thereafter, if the animal spends more time in the chamber associated with the stimulus, a conditioned place preference has developed. If the animal spends less time in the chamber associated with the stimulus, a conditioned place aversion has developed (Cunningham et al. 2003, Cunningham et al. 2006, Cunningham et al. 2011). Depending on the aversive stimulus, the test should be categorised as either “mild” or “moderate”.

5.8.2. Intra Venous Self-Administration

This test is used to evaluate rewarding and abusive effects of drugs. The apparatus is an operant chamber, equipped with two levers, dipper cups, a stimulus light and microliter syringe pumps. First, the mouse is habituated to the operant chamber and trained to press the lever to get a reward, for 1 h per day, for up to one week. Then, the animal undergoes sur-

gery to insert an intravenous catheter subcutaneously from the back, through the jugular vein into the right atrium. After a few days of recuperation time, the test trial is performed. The animal is placed in the operant chamber and the catheter is connected to the computer-operated infusion pump. Pressing a lever results in a drug infusion paired with the illumination of the stimulus light, which stays on for a few more seconds during the out-time period. The mouse is allowed to self-administer for three consecutive sessions of 2 h. The number of infusions as well as the total intake of the administered drug are counted (Kmiotek et al. 2012). We suggest categorising this test as either “moderate” or “severe”, since animals first undergo a surgical procedure and then are fixated due to the catheter system.

5.8.3. Oral Self-Administration

This test is used to evaluate rewarding and abusive effects of drugs. A two-bottle choice paradigm can be used as described in 5.2.4.

5.8.4. Scoring Withdrawal Symptoms

The scoring of withdrawal symptoms is done by observing the mice after drug treatment (Crawley 2000). The test should be categorised as either “mild” or “moderate”, since drug withdrawal can be a painfull procedure (National Research Council 2009) and procedures might vary in trial repetitions.

6. Suggested Severity Assessment

The following table gives a short overview of the analysed and discussed tests.

Table 4: Overview of behavioural tests

Test	Average Suggested Severity					Page
	no burden	mild	moderate	severe	non recovery	
Acoustic Prepulse Inhibition		X	X			66
Active Avoidance		X	X			45
Air Puff Flinch Response		X	X			32
Analysis of Feeding Behaviour	X					53-54
Acoustic Startle Response		X	X			31
Auditory Acuity		X				31
Barnes Maze	X	X				39-41
Basket Test	X					19
Beam Walking Test/Balance Beam Test		X				19-21
Behavioural Sensitisation	X					66
Cafeteria Diet	X	X	X			54-55
Cheese Board Test	X					42
Circadian Wheel Running	X					16
Circling	X	X	X			12
Conditioned Fear			X	X		45-49
Conditioned Place Preference/Aversion		X	X			70
Conditioned Taste Aversion		X	X			49
Conditioned Taste Preference			X	X		56
Conditioning of Eyeblink		X				53
Courtship and Copulation	X					57
Cylinder Test	X					12-13
Delayed Matching-to-Position and Delayed Non-Matching-to-Position Performance (DMTP/DNMTP)	X					42
Elevated Platform		X				62
Elevated Plus-Maze	X					62
Elevated Zero-Maze	X					62
Emergence Test	X					62
Fear/Defence Test Battery		X				63
Fear-Potentiated Startle			X	X		63
Five-Choice Serial Reaction Time Attentional Task	X					52
Flinch Test			X			33
Foot Fault Test	X	X	X			13-14

Test	Average Suggested Severity					Page
	no burden	mild	moderate	severe	non recovery	
Footprint Pattern/CatWalk™ XT	X	X				14-15
Forced Swim Test		X	X	X		68-69
Formalin Test		X	X			33-34
Four Plate Test		X				63
Grip Strenght	X	X				23-25
Hole Poke Exploratory Test		X				16-18
Hot Plate			X			34-36
Hyperlocomotion	X	X				67
Immobilisation/Restraint			X	X		67-68
Intra Venous Self-Administration			X	X		70-71
Isolation Induced Fighting	X	X	X			60
Latent Inhibition		X				66-67
Learned Helplessness				X		69
Learned Safety/Fear Inhibition			X	X		63-64
Light-Dark Test	X					28-30
Marble Burying	X					65
Morris Water Maze		X	X	X		42-45
Nesting (Nest Building)	X					57-58
Nestlet Shredding	X					66
Novelty Suppressed Feeding		X				64
Object Recognition Test	X					52
Olfactory Acuity	X					27
Olfactory Discrimination	X	X				27-28
Open Field Test	X	X				18-19, 62
Oral Self-Administration	X					71
Parenting	X					58
Passive Avoidance		X	X			49
Plantar (Hargreaves) Test		X	X			36
Postweaning Social Isolation	X	X	X			69
Predator Odour		X				64
Pup Retrieval	X	X				59
Pup Ultrasonic Vocalization	X	X				59
Radial Maze		X				49-50
Randall-Selitto Test/Paw-Pressure Test		X	X			36-37
Resident-Intruder Test		X	X			61
Restricted Daily Access		X	X			56
Rotarod Test		X				21-22, 53
Rotometer		X				15-16

Test	Average Suggested Severity					Page
	no burden	mild	moderate	severe	non recovery	
Round-Robin Approach	X	X	X			61
Schedule-Induced Operant Tasks	X					52-53
Scoring Withdrawal Symptoms		X	X			71
Social Deprivation/Social Isolation	X	X	X			69
Social Dominance		X				61
Social Interaction/Recognition Test	X	X	X			50-52
Social Transmission of Food Preference Test	X					59-60
Specialised Diets	X	X	X			56
Staircase Test	X					23
Standard Opponent	X	X	X			60
Step-Down Avoidance		X				50
Sucrose Preference Tests		X				70
T-Maze/Y-Maze Delayed Alternation	X	X				45
T/Y-Maze	X					32
Tail Flick Test		X	X			37-38
Tail Immersion/Hot-Water Tail Flick/Cold-Water Tail Flick Test		X				38-39
Tail Suspension Test			X	X		70
Taste Acuity	X					31
Three Chamber Social Approach Task		X				58
Treadmill Test	X	X	X	X		16
Twenty-Four Hour Consumption	X					56-57
Vertical Pole Test		X				23
Visible Burrow System		X				65
Visual Cliff Assay	X					30
Visual Water Box		X	X			31
Vogel Conflict Test		X	X			65
Von Frey Test		X				32
White Noise		X	X			70
Wire Hang Test		X				25-26
Writhing Test		X	X	X		39
Y-Maze/T-Maze Avoidance		X				50

7. Summary and Conclusion

Nowadays a wide range of literature regarding behavioural tests is available. Therefore, an immense number of different test protocols exists, proceeding behavioural tests with different research methods. An incomplete list of more than 100 behavioural tests is described, not including unpublished tests developed and provided by laboratories on their own. This thesis aimed to review publications describing phenotyping procedures in mice with emphasis on details of the experimental setup which could imply a variation of the burden of the tested animals. According to the information given in these papers a range or a certain grade of severity is suggested in this thesis for every category of phenotyping experiments.
As the findings of this thesis show, research does not use standardised testing protocols for each procedure. This does not only lead to the problem of comparing data and results of different scientific studies but also to the problem that evaluating the exact impact of different parameters on the animal's welfare is very difficult. According to the Directive of the European Parliament, every procedure has to be evaluated individually using the annex VIII. It is important to note that the severity should be classified due to the extent of pain, suffering, fear or permanent damage which the animal is likely to experience during the procedure. Furthermore, the directive mentions some examples of the different severity classifications. A list, with examples of behavioural tests is published by the Austrian Ministry of Education, Science and Research as recommendation for severity classification (https://www.bmbwf.gv.at/Themen/Forschung/Forschung-in-%C3%96sterreich/Services/Tierversuche.html). This list in based on defined experimental parameters. Therefore, the review of this thesis shows that there is a number of varying parameters in most of the tests which could influence the animal's welfare differently. It is plausible that genetically modified animals might have another level of stress, anxiety and fear, which influences the animal's welfare as well. Finally, it is noticeable that the severity assessment of different tests would be more exact by not only using the categories “mild”, “moderate”, “severe” and “non recovery”, but maybe assessing severity by the use of a more precise severity-scale reaching from 1-10.

8. Zusammenfassung

Heutzutage ist eine große Anzahl an Literatur zum Thema Verhaltenstests verfügbar. Deswegen gibt es eine immense Anzahl unterschiedlicher Testprotokolle, welche Verhaltenstests mit unterschiedlichsten Methoden und genutzten Parametern beschreiben. In dieser Diplomarbeit wird eine unvollständige Liste von mehr als 100 Verhaltenstests beschrieben, welche die von Laboratorien selbst entwickelten und unveröffentlichten Tests nicht berücksichtigt. Das Ziel dieser Diplomarbeit war eine Übersicht über die bekanntesten Verhaltenstests zu erstellen und die Unterschiede in den das Tierwohl betreffenden Einflussfaktoren in den verschiedenen Studien zu benennen. Zusätzlich wird eine Einstufung der Schweregrade der jeweiligen Verhaltenstests bei Labormäusen aufgrund der Beschreibungen in der entsprechenden wissenschaftlichen Literatur vorgeschlagen.

Wie im Laufe dieser Diplomarbeit aufgezeigt wird, werden in der Forschung kaum standardisierte Testprotokolle für die einzelnen Verfahren verwendet. Dies führt nicht nur zu dem Problem des Vergleichs von Daten und Ergebnissen verschiedener wissenschaftlicher Hypothesen, sondern auch zu dem Problem, dass eine Bewertung der genauen Auswirkung verschiedener Parameter auf das Wohlergehen des Tieres nahezu unmöglich ist. Nach der Richtlinie des Europäischen Parlaments muss jedes Verfahren anhand des Anhangs VIII einzeln bewertet werden. Dabei wird der Schweregrad aufgrund des Ausmaßes von Schmerzen, Leiden, Angst oder bleibenden Schäden, die das Tier während des Verfahrens wahrscheinlich erleidet, festgelegt. Darüber hinaus werden in der Richtlinie einige Beispiele für die verschiedenen Klassifikationen der Schweregrade genannt. Die Richtlinie beschreibt aber keine genauen Parameter dieser Tests. Es gibt zusätzlich eine Stellungnahme zur Belastungseinstufung in Verhaltensexperimenten des Österreichischen Bundesministerium für Bildung, Wissenschaft und Forschung (https://www.bmbwf.gv.at/Themen/Forschung/Forschung-in-%C3%96sterreich/Services/Tierversuche.html). Darin werden die Einstufungen nach definierten experimentellen Parametern durchgeführt, während sie in der vorliegenden Arbeit anhand der Beschreibungen in den veröffentlichten wissenschaftlichen Studien vorgeschlagen wurden. Die Ergebnisse dieser Arbeit zeigen, dass weitere Studien erforderlich sind, um die Parameter und deren Einfluss auf das Wohlergehen der genutzten Labortiere zu bewerten. Außerdem zeigen eventuell genetisch veränderte Mäuse ein anderes Empfinden in Hinblick auf Stress, Angst und Furcht, was sich damit auch auf das Wohlergehen des Tieres auswirken könnte. Abschließend ist zu sagen, dass die Beurteilung des Schweregrads der verschiedenen Tests genauer wäre,

wenn nicht nur die Kategorien „gering“, „mittel“, „schwer“ und „nicht wiederhergestellt“ verwendet würden, sondern der Schweregrad mithilfe einer genaueren Skala bewertet würde, welche z.B. von 1-10 reicht.

9. References

References

Aartsma-Rus A, van Putten M. 2014. Assessing functional performance in the mdx mouse model. Journal of visualized experiments : JoVE, (85). DOI 10.3791/51303.

Abraira VE, Ginty DD. 2013. The sensory neurons of touch. Neuron, 79 (4): 618–639. DOI 10.1016/j.neuron.2013.07.051.

Alamri FF, Shoyaib AA, Biggers A, Jayaraman S, Guindon J, Karamyan VT. 2018. Applicability of the grip strength and automated von Frey tactile sensitivity tests in the mouse photothrombotic model of stroke. Behavioural Brain Research, 336: 250–255. DOI 10.1016/j.bbr.2017.09.008.

Alexandra Bernardo, Meghan McCord, Aron M. Troen, John D. Allison, Michael P. McDonald. 2007. Impaired spatial memory in APP-overexpressing mice on a homocysteinemia-inducing diet. Neurobiology of aging, 28 (8): 1195–1205. DOI 10.1016/j.neurobiolaging.2006.05.035.

Almond SL, Fradley RL, Armstrong EJ, Heavens RB, Rutter AR, Newman RJ, Chiu CS, Konno R, Hutson PH, Brandon NJ. 2006. Behavioral and biochemical characterization of a mutant mouse strain lacking D-amino acid oxidase activity and its implications for schizophrenia. Molecular and cellular neurosciences, 32 (4): 324–334. DOI 10.1016/j.mcn.2006.05.003.

Angoa-Pérez M, Kane MJ, Briggs DI, Francescutti DM, Kuhn DM. 2013. Marble burying and nestlet shredding as tests of repetitive, compulsive-like behaviors in mice. Journal of visualized experiments : JoVE, (82): 50978. DOI 10.3791/50978.

Anisman H, Merali Z. 2001. Rodent Models of Depression: Learned Helplessness Induced in Mice. Current Protocols in Neuroscience, 14 (1): 8.10C.1-8.10C.15. DOI 10.1002/0471142301.ns0810cs14.

Anthony A, Ackerman E, Lloyd JA. 1959. Noise Stress in Laboratory Rodents. I. Behavioral and Endocrine Response of Mice, Rats, and Guinea Pigs. The Journal of the Acoustical Society of America, 31 (11): 1430–1437. DOI 10.1121/1.1907645.

Antipova V, Wree A, Holzmann C, Mann T, Palomero-Gallagher N, Zilles K, Schmitt O, Hawlitschka A. 2018. Unilateral Botulinum Neurotoxin-A Injection into the Striatum of C57BL/6 Mice Leads to a Different Motor Behavior Compared with Rats. Toxins, 10 (7). DOI 10.3390/toxins10070295.

Ardayfio P, Kim K-S. 2006. Anxiogenic-like effect of chronic corticosterone in the light-dark emergence task in mice. Behavioral neuroscience, 120 (2): 249–256. DOI 10.1037/0735-7044.120.2.249.

Arenas MC, Daza-Losada M, Vidal-Infer A, Aguilar MA, Miñarro J, Rodríguez-Arias M. 2014. Capacity of novelty-induced locomotor activity and the hole-board test to predict sensitivity to the conditioned rewarding effects of cocaine. Physiology & Behavior, 133: 152–160. DOI 10.1016/j.physbeh.2014.05.028.

Arp JM, Ter Horst JP, Loi M, den Blaauwen J, Bangert E, Fernández G, Joëls M, Oitzl MS, Krugers HJ. 2016. Blocking glucocorticoid receptors at adolescent age prevents enhanced freezing between repeated cue-exposures after conditioned fear in adult mice raised under chronic early life stress. Neurobiology of learning and memory, 133: 30–38. DOI 10.1016/j.nlm.2016.05.009.

Azzara AV. 2004. Genetic and biobehavioral analysis of the controls of meal size in mice. Appetite, 42 (1): 107–109. DOI 10.1016/j.appet.2003.08.012.

Bailey CJ, Flatt PR, Kwasowski P, Powell CJ, Marks V. 1986. Immunoreactive gastric inhibitory polypeptide and K cell hyperplasia in obese hyperglycaemic (ob/ob) mice fed high fat and high carbohydrate cafeteria diets. Acta endocrinologica, 112 (2): 224–229. DOI 10.1530/acta.0.1120224.

Bailoo JD, Murphy E, Boada-Saña M, Varholick JA, Hintze S, Baussière C, Hahn KC, Göpfert C, Palme R, Voelkl B, Würbel H. 2018. Effects of Cage Enrichment on Behavior, Welfare and Outcome Variability in Female Mice. Frontiers in behavioral neuroscience, 12: 232. DOI 10.3389/fnbeh.2018.00232.

Baird AL, Meldrum A, Dunnett SB. 2001. The staircase test of skilled reaching in mice. Brain Research Bulletin, 54 (2): 243–250. DOI 10.1016/S0361-9230(00)00457-3.

Bannon AW, Malmberg AB. 2007. Models of nociception: hot-plate, tail-flick, and formalin tests in rodents. Current Protocols in Neuroscience, Chapter 8: Unit 8.9. DOI 10.1002/0471142301.ns0809s41.

Barnes CA. 1979. Memory deficits associated with senescence: a neurophysiological and behavioral study in the rat. Journal of comparative and physiological psychology, 93 (1): 74–104. DOI 10.1037/h0077579.

Baron SP, Meltzer LT. 2001. Mouse strains differ under a simple schedule of operant learning. Behavioural Brain Research, 118 (2): 143–152. DOI 10.1016/S0166-4328(00)00322-3.

Barreto G, Huang T-T, Giffard RG. 2010. Age-related defects in sensorimotor activity, spatial learning, and memory in C57BL/6 mice. Journal of neurosurgical anesthesiology, 22 (3): 214–219. DOI 10.1097/ANA.0b013e3181d56c98.

Baskin YK, Dietrich WD, Green EJ. 2003. Two effective behavioral tasks for evaluating sensorimotor dysfunction following traumatic brain injury in mice. Journal of neuroscience methods, 129 (1): 87–93. DOI 10.1016/S0165-0270(03)00212-7.

Batka RJ, Brown TJ, Mcmillan KP, Meadows RM, Jones KJ, Haulcomb MM. 2014. The Need for Speed in Rodent Locomotion Analyses. The Anatomical Record, 297 (10): 1839–1864. DOI 10.1002/ar.22955.

Bays TB, Lightfoot T, Mayer J, Hrsg. 2006. Exotic Pet Behavior. Saint Louis: W.B. Saunders.

Bean NJ, Nyby J, Kerchner M, Dahinden Z. 1986. Hormonal regulation of chemosignal-stimulated precopulatory behaviors in male housemice (Mus musculus). Hormones and behavior, 20 (4): 390–404. DOI 10.1016/0018-506X(86)90002-4.

Bekkevold CM, Robertson KL, Reinhard MK, Battles AH, Rowland NE. 2013. Dehydration parameters and standards for laboratory mice. Journal of the American Association for Laboratory Animal Science : JAALAS, 52 (3): 233–239. https://pubmed.ncbi.nlm.nih.gov/23849404.

Bercik P, Verdú EF, Foster JA, Lu J, Scharringa A, Kean I, Wang L, Blennerhassett P, Collins SM. 2009. Role of gut-brain axis in persistent abnormal feeding behavior in mice following eradication of Helicobacter pylori infection. American Journal of Physiology-Regulatory, Integrative and Comparative Physiology, 296 (3): R587-R594. DOI 10.1152/ajpregu.90752.2008.

Bernardo A, Harrison FE, McCord M, Zhao J, Bruchey A, Davies SS, Jackson Roberts L, Mathews PM, Matsuoka Y, Ariga T, Yu RK, Thompson R, McDonald MP. 2009. Elimination of GD3 synthase improves memory and reduces amyloid-β plaque load in transgenic mice. Neurobiology of aging, 30 (11): 1777–1791. DOI 10.1016/j.neurobiolaging.2007.12.022.

Bernardo A, McCord M, Troen AM, Allison JD, McDonald MP. 2007. Impaired spatial memory in APP-overexpressing mice on a homocysteinemia-inducing diet. Neurobiology of aging, 28 (8): 1195–1205. DOI 10.1016/j.neurobiolaging.2006.05.035.

Berry A, Bellisario V, Capoccia S, Tirassa P, Calza A, Alleva E, Cirulli F. 2012. Social deprivation stress is a triggering factor for the emergence of anxiety- and depression-like behaviours and leads to reduced brain BDNF levels in C57BL/6J mice. Psychoneuroendocrinology, 37 (6): 762–772. DOI 10.1016/j.psyneuen.2011.09.007.

Bielsky IF, Hu S-B, Szegda KL, Westphal H, Young LJ. 2004. Profound impairment in social recognition and reduction in anxiety-like behavior in vasopressin V1a receptor knockout mice. Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology, 29 (3): 483–493. DOI 10.1038/sj.npp.1300360.

Blanchard, Robert, Parmigiani, Stefano, Agullana, Rey, M. Weiss, Scott, Blanchard, Dixie. 1995. Behaviors of Swiss-Webster and C57/BL/6N sin mice in a fear/defense test battery. Aggressive Behavior, 21. DOI 10.1002/1098-2337(1995)21:1<21::AID-AB2480210105>3.0.CO;2-0.

Blanchard RJ, Griebel G, Henrie JA, Blanchard DC. 1997. Differentiation of anxiolytic and panicolytic drugs by effects on rat and mouse defense test batteries. Neuroscience & Biobehavioral Reviews, 21 (6): 783–789. DOI 10.1016/s0149-7634(96)00062-0.

Blanchard, Robert J., Parmigiani, Stefano, Bjornson, Christopher, Masuda, Camlyn, Weiss, Scott M., Caroline Blanchard, D. 1995. Antipredator behavior of Swiss-Webster mice in a visible burrow system. Aggressive Behavior, 21 (2): 123–136. DOI 10.1002/1098-2337(1995)21:2<123::AID-AB2480210205>3.0.CO;2-E.

Bockamp E, Maringer M, Spangenberg C, Fees S, Fraser S, Eshkind L, Oesch F, Zabel B. 2002. Of mice and models: improved animal models for biomedical research. Physiological genomics, 11 (3): 115–132. DOI 10.1152/physiolgenomics.00067.2002.

Bodden C, Siestrup S, Palme R, Kaiser S, Sachser N, Richter SH. 2018. Evidence-based severity assessment: Impact of repeated versus single open-field testing on welfare in C57BL/6J mice. Behavioural Brain Research, 336: 261–268. DOI 10.1016/j.bbr.2017.08.029.

Bohlen M, Cameron A, Metten P, Crabbe JC, Wahlsten D. 2009. Calibration of rotational acceleration for the rotarod test of rodent motor coordination. Journal of neuroscience methods, 178 (1): 10–14. DOI 10.1016/j.jneumeth.2008.11.001.

Bohn LM, Lefkowitz RJ, Caron MG. 2002. Differential mechanisms of morphine antinociceptive tolerance revealed in (beta)arrestin-2 knock-out mice. The Journal of neuroscience : the official journal of the Society for Neuroscience, 22 (23): 10494–10500.

Boissier JR, Simon P. 1960. L'utilisation du test de la traction (Test de Julou-Courvoisier) pour l'étude des psycholeptiques. Thérapie, (15): 1170–1174.

Borba Filho GL, Zenki KC, Kalinine E, Baggio S, Pettenuzzo L, Zimmer ER, Weis SN, Calcagnotto ME, Onofre de Souza D. 2015. A New Device for Step-Down Inhibitory Avoidance Task—Effects of Low and High Frequency in a Novel Device for Passive Inhibitory Avoidance Task That Avoids Bioimpedance Variations. PloS one, 10 (2): e0116000. DOI 10.1371/journal.pone.0116000.

Boulis NM, Davis M. 1989. Footshock-induced sensitization of electrically elicited startle reflexes. Behavioral neuroscience, 103 (3): 504–508. DOI 10.1037//0735-7044.103.3.504.

Boulton, Alan A., Baker, Glen B., Martin-Iverson, Mathew T., Hrsg. 1991. Animal Models in Psychiatry, II. Humana Press, Totowa, NJ.

Bourin M, Hascoet M. 2003. The mouse light/dark box test. European journal of pharmacology, 463 (1-3): 55–65. DOI 10.1016/s0014-2999(03)01274-3.

Bourin M, Petit-Demoulière B, Nic Dhonnchadha B, Hascöet M. 2007. Animal models of anxiety in mice. Fundamental & Clinical Pharmacology, 21 (6): 567–574. DOI 10.1111/j.1472-8206.2007.00526.x.

Bovet D, Bovet-Nitti F, Oliverio A. 1969. Genetic aspects of learning and memory in mice. Science (New York, N.Y.), 163 (3863): 139–149. DOI 10.1126/science.163.3863.139.

Bowers JM, Alexander BK. 1967. Mice: individual recognition by olfactory cues. Science (New York, N.Y.), 158 (3805): 1208–1210. DOI 10.1126/science.158.3805.1208.

Brain P, Poole A. 1974. Some studies on the use of "standard opponents" in intermale aggression testing in TT albino mice. Behaviour, 50 (1): 100–110. DOI 10.1163/156853974X00057.

Brain PF, Nowell NW. 1970. Activity and defaecation related to aggressiveness and adrenal stress response in adult male laboratory mice. Physiology & Behavior, 5 (3): 259–261. DOI 10.1016/0031-9384(70)90095-8.

Briese E, Cabanac M. 1991. Stress hyperthermia: Physiological arguments that it is a fever. Physiology & Behavior, 49 (6): 1153–1157. DOI 10.1016/0031-9384(91)90343-M.

Bromley-Brits K, Deng Y, Song W, Hrsg. 2011. Morris water maze test for learning and memory deficits in Alzheimer's disease model mice.

Brown JR, Ye H, Bronson RT, Dikkes P, Greenberg ME. 1996. A Defect in Nurturing in Mice Lacking the Immediate Early Gene fosB. Cell, 86 (2): 297–309. DOI 10.1016/S0092-8674(00)80101-4.

Bundesministerium für Bildung, Wissenschaft und Forschung. Tierversuchsstatistiken Österreich.

Buselmaier W, Vierling T, Balzereit W, Schwegler H. 1981. Genetic analysis of avoidance learning by means of different psychological testing systems with inbred mice as model organisms. Psychological Research, 43 (3): 317–333. DOI 10.1007/BF00308455.

Caballero-Garrido E, Pena-Philippides JC, Galochkina Z, Erhardt E, Roitbak T. 2017. Characterization of long-term gait deficits in mouse dMCAO, using the CatWalk system. Behavioural Brain Research, 331: 282–296. DOI 10.1016/j.bbr.2017.05.042.

Cabe PA, Tilson HA, Mitchell CL, Dennis R. 1978. A simple recording grip strength device. Pharmacology, biochemistry, and behavior, 8 (1): 101–102. DOI 10.1016/0091-3057(78)90131-4.

Can A, Dao DT, Arad M, Terrillion CE, Piantadosi SC, Gould TD. 2012. The mouse forced swim test. Journal of visualized experiments : JoVE, (59): 3638-e3638. DOI 10.3791/3638.

Can A, Dao DT, Terrillion CE, Piantadosi SC, Bhat S, Gould TD. 2012. The tail suspension test. Journal of visualized experiments : JoVE, (59): e3769-e3769. DOI 10.3791/3769.

Capoccia S, Maccarinelli F, Buffoli B, Rodella LF, Cremona O, Arosio P, Cirulli F. 2015. Behavioral characterization of mouse models of neuroferritinopathy. PloS one, 10 (2): e0118990. DOI 10.1371/journal.pone.0118990.

Carey LM, Gutierrez T, Deng L, Lee W-H, Mackie K, Hohmann AG. 2017. Inflammatory and Neuropathic Nociception is Preserved in GPR55 Knockout Mice. Scientific Reports, 7 (1): 944. DOI 10.1038/s41598-017-01062-2.

Carter M, Shieh J, Hrsg. 2015. Guide to Research Techniques in Neuroscience (Second Edition). San Diego: Academic Press.

Castro B, Kuang S. 2017. Evaluation of Muscle Performance in Mice by Treadmill Exhaustion Test and Whole-limb Grip Strength Assay. Bio-protocol, 7 (8): e2237. DOI 10.21769/BioProtoc.2237.

Chambers KC. 2018. Conditioned taste aversions. World journal of otorhinolaryngology - head and neck surgery, 4 (1): 92–100. DOI 10.1016/j.wjorl.2018.02.003.

Chaudhury D, Colwell CS. 2002. Circadian modulation of learning and memory in fear-conditioned mice. Behavioural Brain Research, 133 (1): 95–108. DOI 10.1016/s0166-4328(01)00471-5.

Chu X, Zhou Y, Hu Z, Lou J, Song W, Li J, Liang X, Chen C, Wang S, Yang B, Chen L, Zhang X, Song J, Dong Y, Chen S, He L, Xie Q, Chen X, Li W. 2016. 24-hour-restraint stress induces long-term depressive-like phenotypes in mice. Scientific Reports, 6: 32935 EP -. DOI 10.1038/srep32935.

Clarkson AN, Huang BS, MacIsaac SE, Mody I, Carmichael ST. 2010. Reducing excessive GABA-mediated tonic inhibition promotes functional recovery after stroke. Nature, 468 (7321): 305–309. DOI 10.1038/nature09511.

Clarkson JM, Dwyer DM, Flecknell PA, Leach MC, Rowe C. 2018. Handling method alters the hedonic value of reward in laboratory mice. Scientific Reports, 8 (1): 2448. DOI 10.1038/s41598-018-20716-3 (accessed Apr 14, 2019).

Clement JG. 1993. Experimentally induced mortality following repeated measurement of rectal temperature in mice. Laboratory Animal Science, 43 (4): 381–382. https://www.scopus.com/inward/record.uri?eid=2-s2.0-0027424114&partnerID=40&md5=342cc55e742fa9a43e67f616e62088c0.

Clement JG, Mills P, Brockway B. 1989. Use of telemetry to record body temperature and activity in mice. Journal of pharmacological methods, 21 (2): 129–140. DOI 10.1016/0160-5402(89)90031-4.

Colla ARS, Rosa JM, Cunha MP, Rodrigues ALS. 2015. Anxiolytic-like effects of ursolic acid in mice. European journal of pharmacology, 758: 171–176. DOI 10.1016/j.ejphar.2015.03.077.

Cook MN, Crounse M, Flaherty L. 2002. Anxiety in the Elevated Zero-Maze Is Augmented in Mice After Repeated Daily Exposure. Behavior Genetics, 32 (2): 113–118. DOI 10.1023/A:1015249706579.

Costa RM, Honjo T, Silva AJ. 2003. Learning and memory deficits in Notch mutant mice. Current biology : CB, 13 (15): 1348–1354. DOI 10.1016/s0960-9822(03)00492-5.

Costa RM, Yang T, Huynh DP, Pulst SM, Viskochil DH, Silva AJ, Brannan CI. 2001. Learning deficits, but normal development and tumor predisposition, in mice lacking exon 23a of Nf1. Nature Genetics, 27 (4): 399–405. DOI 10.1038/86898.

Crawley J, Goodwin FK. 1980. Preliminary report of a simple animal behavior model for the anxiolytic effects of benzodiazepines. Pharmacology, biochemistry, and behavior, 13 (2): 167–170. DOI 10.1016/0091-3057(80)90067-2.

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

Crawley JN. 2000. What's wrong with my mouse? Behavioral phenotyping of transgenic and knockout mice. Hoboken, NJ, USA: John Wiley & Sons, Inc; Wiley-Liss, 329.

Crusio WE, Schwegler H. 2005. Learning spatial orientation tasks in the radial-maze and structural variation in the hippocampus in inbred mice. Behavioral and Brain Functions, 1 (1): 3. DOI 10.1186/1744-9081-1-3.

Cui F, Liu H, Zou Z, Li H. 2015. Metabolic responses to water deprivation in C57BL/6J mice using a proton nuclear magnetic resonance-based metabonomics approach. RSC Advances, 5 (98): 80142–80149. DOI 10.1039/C5RA12650A.

Cunningham CL, Ferree NK, Howard MA. 2003. Apparatus bias and place conditioning with ethanol in mice. Psychopharmacology, 170 (4): 409–422. DOI 10.1007/s00213-003-1559-y.

Cunningham CL, Gremel CM, Groblewski PA. 2006. Drug-induced conditioned place preference and aversion in mice. Nature protocols, 1 (4): 1662–1670. DOI 10.1038/nprot.2006.279.

Cunningham CL, Groblewski PA, Voorhees CM, Olmstead MC. 2011. Place Conditioning. First ed. Totowa, NJ: Springer Science+Business Media LLC.

Czyzyk TA. 2013. Assessment of stress-independent binge-like eating behavior in mice. Neuromethods, 74 (DOI: 10.1007/978-1-62703-104-2-5). DOI 10.1007/978-1-62703-104-2-5.

Dalla Costa E, Minero M, Lebelt D, Stucke D, Canali E, Leach MC. 2014. Development of the Horse Grimace Scale (HGS) as a pain assessment tool in horses undergoing routine castration. PloS one, 9 (3): e92281. DOI 10.1371/journal.pone.0092281.

Deacon RMJ. 2006. Assessing nest building in mice. Nature protocols, 1 (3): 1117–1119. DOI 10.1038/nprot.2006.170.

Deacon RMJ. 2013. Measuring motor coordination in mice. Journal of visualized experiments : JoVE, (75): e2609. DOI 10.3791/2609.

Deacon RMJ, Rawlins JNP. 2006. T-maze alternation in the rodent. Nature protocols, 1 (1): 7–12. DOI 10.1038/nprot.2006.2.

Delcasso S, Jeantet Y, Cho YH. 2007. A new test for long-term spatial memory using an operant chamber in mice. Behavioural Brain Research, 178 (2): 200–207. DOI 10.1016/j.bbr.2006.12.019.

Delini-Stula A, Vassout A. 1981. The Effects of Antidepressants on Aggressiveness Induced by Social Deprivation in Mice. Pharmacology, biochemistry, and behavior, 14: 33–41. DOI 10.1016/S0091-3057(81)80008-1.

Deuis JR, Dvorakova LS, Vetter I. 2017. Methods Used to Evaluate Pain Behaviors in Rodents. Frontiers in molecular neuroscience, 10: 284. DOI 10.3389/fnmol.2017.00284.

Disterhoft JF, Weiss C. 2017. 3.19 - Eyeblink Conditioning – A Behavioral Model of Procedural and Declarative Learning. In: Byrne JH, ed. Learning and Memory: A Comprehensive Reference (Second Edition). Oxford: Academic Press, 327–355.

Dogrul A, Gulmez SE, Deveci MS, Gul H, Ossipov MH, Porreca F, Tulunay FC. 2007. The local antinociceptive actions of nonsteroidal antiinflammatory drugs in the mouse radiant heat tail-flick test. Anesthesia and analgesia, 104 (4): 927–935. DOI 10.1213/01.ane.0000258773.46897.34.

Dorman CW, Krug HE, Frizelle SP, Funkenbusch S, Mahowald ML. 2014. A comparison of DigiGait™ and TreadScan™ imaging systems: assessment of pain using gait analysis in murine monoarthritis. Journal of pain research, 7: 25–35. DOI 10.2147/JPR.S52195.

Drel VR, Pacher P, Vareniuk I, Pavlov I, Ilnytska O, Lyzogubov VV, Tibrewala J, Groves JT, Obrosova IG. 2007. A peroxynitrite decomposition catalyst counteracts sensory neuropathy in streptozotocin-diabetic mice. European journal of pharmacology, 569 (1-2): 48–58. DOI 10.1016/j.ejphar.2007.05.055.

Drickamer LC. 2001. Urine marking and social dominance in male house mice (Mus musculus domesticus). Behavioural Processes, 53 (1): 113–120. DOI 10.1016/S0376-6357(00)00152-2.

Dunham NW, Miya TS. 1957. A note on a simple apparatus for detecting neurological deficit in rats and mice. Journal of the American Pharmaceutical Association. American Pharmaceutical Association, 46 (3): 208–209. DOI 10.1002/jps.3030460322.

Eide PK, Berge OG, Tjolsen A, Hole K. 1988. Apparent hyperalgesia in the mouse tail-flick test due to increased tail skin temperature after lesioning of serotonergic pathways. Acta physiologica Scandinavica, 134 (3): 413–420. DOI 10.1111/j.1748-1716.1988.tb08509.x.

Ennaceur A. 2012. Open Space Anxiety Test in Rodents: The Elevated Platform with Steep Slopes. Methods in molecular biology (Clifton, N.J.), 829. DOI 10.1007/978-1-61779-458-2_11.

Enwere E, Shingo T, Gregg C, Fujikawa H, Ohta S, Weiss S. 2004. Aging results in reduced epidermal growth factor receptor signaling, diminished olfactory neurogenesis, and deficits in fine olfactory discrimination. The Journal of neuroscience : the official journal of the Society for Neuroscience, 24 (38): 8354–8365. DOI 10.1523/JNEUROSCI.2751-04.2004.

Ernst L, Kopaczka M, Schulz M, Talbot SR, Struve B, Häger C, Bleich A, Durst M, Jirkof P, Arras M, van Dijk RM, Miljanovic N, Potschka H, Merhof D, Tolba RH. 2020. Semi-automated generation of pictures for the Mouse Grimace Scale: A multi-laboratory analysis (Part 2). Laboratory animals, 54 (1): 92–98. DOI 10.1177/0023677219881664.

Ernst L, Kopaczka M, Schulz M, Talbot SR, Zieglowski L, Meyer M, Bruch S, Merhof D, Tolba RH. 2020. Improvement of the Mouse Grimace Scale set-up for implementing a semi-automated Mouse Grimace Scale scoring (Part 1). Laboratory animals, 54 (1): 83–91. DOI 10.1177/0023677219881655.

European parliament and council of 22 September 2010. 2010. Directive 2010/63/EU of the European Parliament and of the Council of 22 September 2010 on the protection of animals used for scientific purposes.

Evangelista MC, Watanabe R, Leung VSY, Monteiro BP, O'Toole E, Pang DSJ, Steagall PV. 2019. Facial expressions of pain in cats: the development and validation of a Feline Grimace Scale. Scientific Reports, 9 (1): 19128. DOI 10.1038/s41598-019-55693-8.

Fahlstrom A, Yu Q, Ulfhake B. 2011. Behavioral changes in aging female C57BL/6 mice. Neurobiology of aging, 32 (10): 1868–1880. DOI 10.1016/j.neurobiolaging.2009.11.003.

Faizi M, Bader PL, Saw N, Nguyen T-VV, Beraki S, Wyss-Coray T, Longo FM, Shamloo M. 2012. Thy1-hAPP(Lond/Swe+) mouse model of Alzheimer's disease displays broad behavioral deficits in sensorimotor, cognitive and social function. Brain and behavior, 2 (2): 142–154. DOI 10.1002/brb3.41.

Fajrin FA, Susilowati R, Nurrochmad A, Nugroho AE. 2017. The influence of long-term diabetes mellitus (DM) on pain response in mice: in vivo models of painful diabetic neuropathy (PDN). Journal of thee Medical Sciences (Berkala Ilmu Kedokteran), 49 (03): 97–105. DOI 10.19106/JMedSci004903201701.

Fan Y, Shen F, Frenzel T, Zhu W, Ye J, Liu J, Chen Y, Su H, Young WL, Yang G-Y. 2010. Endothelial progenitor cell transplantation improves long-term stroke outcome in mice. Annals of neurology, 67 (4): 488–497. DOI 10.1002/ana.21919.

Ferguson JN, Young LJ, Hearn EF, Matzuk MM, Insel TR, Winslow JT. 2000. Social amnesia in mice lacking the oxytocin gene. Nature Genetics, 25 (3): 284–288. DOI 10.1038/77040.

Fink D, Yau TY, Kolbe T, Rülicke T. 2015. Non-invasive instant genotyping of fluorescently labelled transgenic mice. ALTEX, 32 (3): 222–227. DOI 10.14573/altex.1502181.

Fink G, Hrsg. 2007. Encyclopedia of Stress (Second Edition). New York: Academic Press.

Flores-Montoya MG, Sobin C. 2015. Early chronic lead exposure reduces exploratory activity in young C57BL/6J mice. Journal of applied toxicology : JAT, 35 (7): 759–765. DOI 10.1002/jat.3064.

Fox JG, Hrsg. 2007. The Mouse in biomedical research. Secondnd [upd. and rev.] ed. Amsterdam, New York: Academic Press, 1 online resource (4 volumes (a.

Fox MW. 1965. The visual cliff test for the study of visual depth perception in the mouse. Animal Behaviour, 13 (2): 232-IN3. DOI 10.1016/0003-3472(65)90040-0.

Foy MR, Foy JG. 2017. Conditioned Taste Aversion☆. In: . Reference Module in Neuroscience and Biobehavioral Psychology. : Elsevier.

Gapp K, Jawaid A, Sarkies P, Bohacek J, Pelczar P, Prados J, Farinelli L, Miska E, Mansuy IM. 2014. Implication of sperm RNAs in transgenerational inheritance of the effects of early trauma in mice. Nature neuroscience, 17 (5): 667–669. DOI 10.1038/nn.3695.

Gaskill BN, Karas AZ, Garner JP, Pritchett-Corning KR. 2013. Nest building as an indicator of health and welfare in laboratory mice. Journal of visualized experiments : JoVE, (82): 51012. DOI 10.3791/51012.

Gasparin FRS, Carreño FO, Mewes JM, Gilglioni EH, Pagadigorria CLS, Natali MRM, Utsunomiya KS, Constantin RP, Ouchida AT, Curti C, Gaemers IC, Elferink, Ronald Petrus Johannes Oude, Constantin J, Ishii-Iwamoto EL. 2018. Sex differences in the development of hepatic steatosis in cafeteria diet-induced obesity in young mice. Biochimica et biophysica acta. Molecular basis of disease, 1864 (7): 2495–2509. DOI 10.1016/j.bbadis.2018.04.004.

Gawade SP. 2012. Acetic acid induced painful endogenous infliction in writhing test on mice. Journal of pharmacology & pharmacotherapeutics, 3 (4): 348. DOI 10.4103/0976-500X.103699.

Gerlai R, Thibodeaux H, Palmer JT, van Lookeren Campagne M, van Bruggen N. 2000. Transient focal cerebral ischemia induces sensorimotor deficits in mice. Behavioural Brain Research, 108 (1): 63–71. DOI 10.1016/s0166-4328(99)00130-8.

Ghafari M, Patil SS, Hoger H, Pollak A, Lubec G. 2011. NMDA-complexes linked to spatial memory performance in the Barnes maze in CD1 mice. Behavioural Brain Research, 221 (1): 142–148. DOI 10.1016/j.bbr.2011.02.042.

Gjendal K, Franco NH, Ottesen JL, Sørensen DB, Olsson IAS. 2018. Eye, body or tail? Thermography as a measure of stress in mice. Physiology & Behavior, 196: 135–143. DOI 10.1016/j.physbeh.2018.08.022.

Gjendal K, Ottesen JL, Olsson IAS, Sørensen DB. 2019. Burrowing and nest building activity in mice after exposure to grid floor, isoflurane or ip injections. Physiology & Behavior, 206: 59–66. DOI 10.1016/j.physbeh.2019.02.022.

Glavin GB, Paré WP, Sandbak T, Bakke H-K, Murison R. 1994. Restraint stress in biomedical research: An update. Neuroscience & Biobehavioral Reviews, 18 (2): 223–249. DOI 10.1016/0149-7634(94)90027-2.

Gleerup KB, Andersen PH, Munksgaard L, Forkman B. 2015. Pain evaluation in dairy cattle. Applied Animal Behaviour Science, 171: 25–32. DOI 10.1016/j.applanim.2015.08.023.

Goto K, Kurashima R, Watanabe S. 2010. Delayed matching-to-position performance in C57BL/6N mice. Behavioural Processes, 84 (2): 591–597. DOI 10.1016/j.beproc.2010.02.022.

Gouveia K, Hurst JL. 2013. Reducing Mouse Anxiety during Handling: Effect of Experience with Handling Tunnels. PloS one, 8 (6): e66401. DOI 10.1371/journal.pone.0066401.

Gouveia K, Hurst JL. 2017. Optimising reliability of mouse performance in behavioural testing: the major role of non-aversive handling. Scientific Reports, 7 (1): 44999. DOI 10.1038/srep44999.

Grailhe R, Waeber C, Dulawa SC, Hornung JP, Zhuang X, Brunner D, Geyer MA, Hen R. 1999. Increased exploratory activity and altered response to LSD in mice lacking the 5-HT(5A) receptor. Neuron, 22 (3): 581–591. DOI 10.1016/s0896-6273(00)80712-6.

Greenberg GD, Phillips TJ, Crabbe JC. 2016. Effects of acute alcohol withdrawal on nest building in mice selectively bred for alcohol withdrawal severity. Physiology & Behavior, 165: 257–266. DOI 10.1016/j.physbeh.2016.08.006.

Guesgen MJ, Beausoleil NJ, Leach M, Minot EO, Stewart M, Stafford KJ. 2016. Coding and quantification of a facial expression for pain in lambs. Behavioural Processes, 132: 49–56. DOI 10.1016/j.beproc.2016.09.010.

Guilford BL, Ryals JM, Wright DE. 2011. Phenotypic changes in diabetic neuropathy induced by a high-fat diet in diabetic C57BL/6 mice. Experimental diabetes research, 2011: 848307. DOI 10.1155/2011/848307.

Guzmán YF, Tronson NC, Guedea A, Huh KH, Gao C, Radulovic J. 2009. Social modeling of conditioned fear in mice by non-fearful conspecifics. Behavioural Brain Research, 201 (1): 173–178. DOI 10.1016/j.bbr.2009.02.024.

Hadfield MG, Nugent EA, Mott DEW. 1982. Cocaine increases isolation-induced fighting in mice. Pharmacology, biochemistry, and behavior, 16 (2): 359–360. DOI 10.1016/0091-3057(82)90172-1.

Häger C, Biernot S, Buettner M, Glage S, Keubler LM, Held N, Bleich EM, Otto K, Müller CW, Decker S, Talbot SR, Bleich A. 2017. The Sheep Grimace Scale as an indicator of post-operative distress and pain in laboratory sheep. PloS one, 12 (4): e0175839. DOI 10.1371/journal.pone.0175839.

Hamers FPT, Koopmans GC, Joosten EAJ. 2006. CatWalk-assisted gait analysis in the assessment of spinal cord injury. Journal of neurotrauma, 23 (3-4): 537–548. DOI 10.1089/neu.2006.23.537.

Hampshire V, Robertson S. 2015. Using the facial grimace scale to evaluate rabbit wellness in post-procedural monitoring. Lab animal, 44 (7): 259–260. DOI 10.1038/laban.806.

Hargreaves K, Dubner R, Brown F, Flores C, Joris J. 1988. A new and sensitive method for measuring thermal nociception in cutaneous hyperalgesia. Pain, 32 (1): 77–88. DOI 10.1016/0304-3959(88)90026-7.

Harrison FE, Reiserer RS, Tomarken AJ, McDonald MP. 2006. Spatial and nonspatial escape strategies in the Barnes maze. Learning & memory (Cold Spring Harbor, N.Y.), 13 (6): 809–819. DOI 10.1101/lm.334306.

Harro J. 2018. Animals, anxiety, and anxiety disorders: How to measure anxiety in rodents and why. Behavioural Brain Research, 352: 81–93. DOI 10.1016/j.bbr.2017.10.016.

Hartsock TG. 1982. Ethological Approach to Farm Animal Behavior Research. Journal of Animal Science, 54 (2): 447–449. DOI 10.2527/jas1982.542447x.

Hascoet M, Bourin M. 1998. A new approach to the light/dark test procedure in mice. Pharmacology, biochemistry, and behavior, 60 (3): 645–653. DOI 10.1016/s0091-3057(98)00031-8.

Hascoët M. BM. 2011. The Four-Plate Test in Mice. Mood and Anxiety Related Phenotypes in Mice, Neuromethods, (63): 123–141. DOI 10.1007/978-1-61779-313-4_8.

Hatori M, Vollmers C, Zarrinpar A, DiTacchio L, Bushong EA, Gill S, Leblanc M, Chaix A, Joens M, Fitzpatrick JAJ, Ellisman MH, Panda S. 2012. Time-Restricted Feeding without Reducing Caloric Intake Prevents Metabolic Diseases in Mice Fed a High-Fat Diet. Cell Metabolism, 15 (6): 848–860. DOI 10.1016/j.cmet.2012.04.019.

Heredia L, Torrente M, Colomina MT, Domingo JL. 2014. Assessing anxiety in C57BL/6J mice: a pharmacological characterization of the open-field and light/dark tests. Journal of pharmacological and toxicological methods, 69 (2): 108–114. DOI 10.1016/j.vascn.2013.12.005.

Hetze S, Römer C, Teufelhart C, Meisel A, Engel O. 2012. Gait analysis as a method for assessing neurological outcome in a mouse model of stroke. Journal of neuroscience methods, 206 (1): 7–14. DOI 10.1016/j.jneumeth.2012.02.001.

Heyser CJ, Vishnevetsky D, Berten S. 2013. The effect of cocaine on rotarod performance in male C57BL/6J mice. Physiology & Behavior, 118: 208–211. DOI 10.1016/j.physbeh.2013.05.027.

Higaki A, Mogi M, Iwanami J, Min L-J, Bai H-Y, Shan B-S, Kukida M, Kan-No H, Ikeda S, Higaki J, Horiuchi M. 2018. Predicting outcome of Morris water maze test in vascular dementia mouse model with deep learning. PloS one, 13 (2): e0191708. DOI 10.1371/journal.pone.0191708.

Hofer MA. 2002. Ultrasonic vocalizations in rat and mouse pups. Current protocols in neuroscience / editorial board, Jacqueline N. Crawley … [et al.], Chapter 8 (ISSN: 1934-8576).

Hoffman E, Winder SJ. 2016. A Modified Wire Hanging Apparatus for Small Animal Muscle Function Testing. PLoS currents, 8. DOI 10.1371/currents.md.1e2bec4e78697b7b0ff80ea25a1d38be.

Hohlbaum K, Bert B, Dietze S, Palme R, Fink H, Thöne-Reineke C. 2018. Impact of repeated anesthesia with ketamine and xylazine on the well-being of C57BL/6JRj mice. PloS one, 13 (9): e0203559-e0203559. DOI 10.1371/journal.pone.0203559.

Huang R-R, Hu W, Yin Y-Y, Wang Y-C, Li W-P, Li W-Z. 2015. Chronic restraint stress promotes learning and memory impairment due to enhanced neuronal endoplasmic reticulum stress in the frontal cortex and hippocampus in male mice. International journal of molecular medicine, 35 (2): 553–559. DOI 10.3892/ijmm.2014.2026.

Huang T-N, Yen T-L, Qiu LR, Chuang H-C, Lerch JP, Hsueh Y-P. 2019. Haploinsufficiency of autism causative gene Tbr1 impairs olfactory discrimination and neuronal activation of the olfactory system in mice. Molecular autism, 10: 5. DOI 10.1186/s13229-019-0257-5.

Hunskaar S, Fasmer OB, Hole K. 1985. Formalin test in mice, a useful technique for evaluating mild analgesics. Journal of neuroscience methods, 14 (1): 69–76. DOI 10.1016/0165-0270(85)90116-5.

Hurst JL, West RS. 2010. Taming anxiety in laboratory mice. Nature methods, 7 (10): 825–826. DOI 10.1038/nmeth.1500.

Institut für Arbeitsschutz der Deutschen Gesetzlichen Unfallversicherung. 04.09.2020. https://www.dguv.de/ifa/gestis/gestis-stoffdatenbank/index.jsp (accessed Sep 10, 2020).

Irwin S. 1968. Comprehensive observational assessment: Ia. A systematic, quantitative procedure for assessing the behavioral and physiologic state of the mouse. Psychopharmacologia, 13 (3): 222–257. DOI 10.1007/BF00401402.

Jacobs SA, Huang F, Tsien JZ, Wei W. 2016. Social Recognition Memory Test in Rodents. Bio-protocol, 6 (9): e1804. DOI 10.21769/BioProtoc.1804.

James BM, Li Q, Luo L, Kendrick KM. 2015. Aged neuronal nitric oxide knockout mice show preserved olfactory learning in both social recognition and odor-conditioning tasks. Frontiers in cellular neuroscience, 9: 105. DOI 10.3389/fncel.2015.00105.

Jang HM, Jang S-E, Han MJ, Kim D-H. 2018. Anxiolytic-like effect of Bifidobacterium adolescentis IM38 in mice with or without immobilisation stress. Beneficial microbes, 9 (1): 123–132. DOI 10.3920/BM2016.0226.

Jänicke B, Coper H. 1996. Tests in Rodents for Assessing Sensorimotor Performance During Aging. In: Ferrandez A-M, Teasdale N, eds. Advances in Psychology : Changes in Sensory Motor Behavior in Aging. : North-Holland, 201–233.

Janus C. 2004. Search strategies used by APP transgenic mice during navigation in the Morris water maze. Learning & memory (Cold Spring Harbor, N.Y.), 11 (3): 337–346. DOI 10.1101/lm.70104.

Jensen P. 2009. The Ethology of Domestic Animals: An Introductory Text. CABI.

Jeremy Bentham. 1823. An Introduction to the Principles of Morals and Legislation. Band 2. Second ed. London: W. Pickering.

Jirkof P, Fleischmann T, Cesarovic N, Rettich A, Vogel J, Arras M. 2013. Assessment of postsurgical distress and pain in laboratory mice by nest complexity scoring. Laboratory animals, 47 (3): 153–161. DOI 10.1177/0023677213475603.

Jung E-S, Lee HJ, Sim H-R, Baik J-H. 2013. Cocaine-induced behavioral sensitization in mice: effects of microinjection of dopamine d2 receptor antagonist into the nucleus accumbens. Experimental neurobiology, 22 (3): 224–231. DOI 10.5607/en.2013.22.3.224.

Kalter H. 1954. Preliminary studies on the metabolic factors involved in the prodction of cleft palate in mice. Genetics, (39): 975.

Kamprath K, Wotjak CT. 2004. Nonassociative learning processes determine expression and extinction of conditioned fear in mice. Learning & memory (Cold Spring Harbor, N.Y.), 11 (6): 770–786. DOI 10.1101/lm.86104.

Kedia S, Chattarji S. 2014. Marble burying as a test of the delayed anxiogenic effects of acute immobilisation stress in mice. Journal of neuroscience methods, 233: 150–154. DOI 10.1016/j.jneumeth.2014.06.012.

Keyhanfar F, Shamsi Meymandi M, Sepehri G, Rastegaryanzadeh R, Heravi G. 2013. Evaluation of Antinociceptive Effect of Pregabalin in Mice and its Combination with Tramadol using Tail Flick Test. Iranian journal of pharmaceutical research : IJPR, 12 (3): 483–493.

Khandve B. 2013. Deranged Spermatogenesis of Adult Swiss Albino Mice as Effect of Immobilization Stress -Histological Study. IOSR Journal of Pharmacy (IOSRPHR), 3 (2): 7–10. DOI 10.9790/3013-3210710.

Kim H, Kang H, Heo RW, Jeon BT, Yi C-O, Shin HJ, Kim J, Jeong S-Y, Kwak W, Kim W-H, Kang SS, Roh GS. 2016. Caloric restriction improves diabetes-induced cognitive deficits by attenuating neurogranin-associated calcium signaling in high-fat diet-fed mice. Journal of cerebral blood flow and metabolism : official journal of the International Society of Cerebral Blood Flow and Metabolism, 36 (6): 1098–1110. DOI 10.1177/0271678X15606724.

Kmiotek EK, Baimel C, Gill KJ. 2012. Methods for intravenous self administration in a mouse model. Journal of visualized experiments : JoVE, (70): 3739-e3739. DOI 10.3791/3739.

Kogan JH, Franklandand PW, Silva AJ. 2000. Long-term memory underlying hippocampus-dependent social recognition in mice. Hippocampus, 10 (1): 47–56. DOI 10.1002/(SICI)1098-1063(2000)10:1<47::AID-HIPO5>3.0.CO;2-6.

König B. 2012. The Behaviour of the House Mouse. In: . The Laboratory Mouse. : 367–381.

Koolhaas JM, Coppens CM, Boer SF de, Buwalda B, Meerlo P, Timmermans PJA. 2013. The resident-intruder paradigm: a standardized test for aggression, violence and social stress. Journal of visualized experiments : JoVE, (77): 4367-e4367. DOI 10.3791/4367.

Koski SK, Leino S, Panula P, Rannanpää S, Salminen O. 2020. Genetic lack of histamine upregulates dopamine neurotransmission and alters rotational behavior but not levodopa-induced dyskinesia in a mouse model of Parkinson's disease. Neuroscience Letters, 729: 134932. DOI 10.1016/j.neulet.2020.134932.

Kraeuter A-K, Guest PC, Sarnyai Z. 2019. The Nest Building Test in Mice for Assessment of General Well-Being. Methods in molecular biology (Clifton, N.J.), 1916: 87–91. DOI 10.1007/978-1-4939-8994-2_7.

Kulesskaya N, Voikar V. 2014. Assessment of mouse anxiety-like behavior in the light-dark box and open-field arena: role of equipment and procedure. Physiology & Behavior, 133: 30–38. DOI 10.1016/j.physbeh.2014.05.006.

Kumar A, Garg R, Prakash AK. 2010. Effect of St. John's Wort (Hypericum perforatum) treatment on restraint stress-induced behavioral and biochemical alteration in mice. BMC complementary and alternative medicine, 10: 18. DOI 10.1186/1472-6882-10-18.

Kunkhyen T, Perez E, Bass M, Coyne A, Baum MJ, Cherry JA. 2018. Gonadal hormones, but not sex, affect the acquisition and maintenance of a Go/No-Go odor discrimination task in mice. Hormones and behavior, 100: 12–19. DOI 10.1016/j.yhbeh.2018.02.009.

Kuroda KO, Meaney MJ, Uetani N, Fortin Y, Ponton A, Kato T. 2007. ERK-FosB signaling in dorsal MPOA neurons plays a major role in the initiation of parental behavior in mice. Molecular and cellular neurosciences, 36 (2): 121–131.
DOI 10.1016/j.mcn.2007.05.010.

Labots MM, Laarakker MCM, Schetters DD, Arndt SSS, van Lith, H A Hein. 2018. An improved procedure for integrated behavioral z-scoring illustrated with modified Hole Board behavior of male inbred laboratory mice. Journal of neuroscience methods, 293: 375–388. DOI 10.1016/j.jneumeth.2017.09.003.

Lalonde R. 1987. Motor abnormalities in staggerer mutant mice. Experimental Brain Research, 68 (2): 417–420. DOI 10.1007/BF00248807.

Langford DJ, Bailey AL, Chanda ML, Clarke SE, Drummond TE, Echols S, Glick S, Ingrao J, Klassen-Ross T, Lacroix-Fralish ML, Matsumiya L, Sorge RE, Sotocinal SG, Tabaka JM, Wong D, van den Maagdenberg, Arn M J M, Ferrari MD, Craig KD, Mogil JS. 2010. Coding of facial expressions of pain in the laboratory mouse. Nature methods, 7 (6): 447–449. DOI 10.1038/nmeth.1455.

Laska M, Sievert T. 2016. Behavioral Responses of CD-1 Mice to Six Predator Odor Components. Chemical Senses, 41 (5): 399–406. DOI 10.1093/chemse/bjw015.

Lee JW, Lee EJ, Hong SH, Chung WH, Lee HT, Lee TW, Lee JR, Kim HT, Suh JG, Kim TY, Ryoo ZY. 2001. Circling mouse: possible animal model for deafness. Comparative medicine, 51 (6): 550–554.

Lee SA, Holly KS, Voziyanov V, Villalba SL, Tong R, Grigsby HE, Glasscock E, Szele FG, Vlachos I, Murray TA. 2016. Gradient Index Microlens Implanted in Prefrontal Cortex of

Mouse Does Not Affect Behavioral Test Performance over Time. PloS one, 11 (1): e0146533. DOI 10.1371/journal.pone.0146533.

Leffa DD, Rezin GT, Daumann F, Longaretti LM, Dajori ALF, Gomes LM, Silva MC, Streck EL, Andrade VM de. 2017. Effects of Acerola (Malpighia emarginata DC.) Juice Intake on Brain Energy Metabolism of Mice Fed a Cafeteria Diet. Molecular neurobiology, 54 (2): 954–963. DOI 10.1007/s12035-016-9691-y.

Leffa DD, Valvassori SS, Varela RB, Lopes-Borges J, Daumann F, Longaretti LM, Dajori ALF, Quevedo J, Andrade VM. 2015. Effects of palatable cafeteria diet on cognitive and noncognitive behaviors and brain neurotrophins' levels in mice. Metabolic brain disease, 30 (4): 1073–1082. DOI 10.1007/s11011-015-9682-0.

Leger M, Quiedeville A, Bouet V, Haelewyn B, Boulouard M, Schumann-Bard P, Freret T. 2013. Object recognition test in mice. Nature protocols, 8: 2531 EP -. DOI 10.1038/nprot.2013.155.

Leroy T, Silva M, D'Hooge R, Aerts J-M, Berckmans D. 2009. Automated gait analysis in the open-field test for laboratory mice. Behavior research methods, 41. DOI 10.3758/BRM.41.1.148.

Li P-C, Jiao Y, Ding J, Chen Y-C, Cui Y, Qian C, Yang X-Y, Ju S-H, Yao H-H, Teng G-J. 2015. Cystamine Improves Functional Recovery via Axon Remodeling and Neuroprotection after Stroke in Mice. CNS Neuroscience & Therapeutics, 21 (3): 231–240. DOI 10.1111/cns.12343.

Lichtman AH, Smith FL, Martin BR. 1993. Evidence that the antinociceptive tail-flick response is produced independently from changes in either tail-skin temperature or core temperature. Pain, 55 (3): 283–295.

Lijam N, Paylor R, McDonald MP, Crawley JN, Deng CX, Herrup K, Stevens KE, Maccaferri G, McBain CJ, Sussman DJ, Wynshaw-Boris A. 1997. Social interaction and sensorimotor gating abnormalities in mice lacking Dvl1. Cell, 90 (5): 895–905. DOI 10.1016/s0092-8674(00)80354-2.

Lindzey G, Delbert DT. 1970. Contributions to behavior-genetic analysis; the mouse as a prototype [by] Gardner Lindzey [and] Delbert D. Thiessen // Contributions to behavior-genetic analysis. The mouse as a prototype. New York: Appleton-Century-Crofts; Appleton, 336.

Liu M-Y, Yin C-Y, Zhu L-J, Zhu X-H, Xu C, Luo C-X, Chen H, Zhu D-Y, Zhou Q-G. 2018. Sucrose preference test for measurement of stress-induced anhedonia in mice. Nature protocols, 13 (7): 1686–1698. DOI 10.1038/s41596-018-0011-z.

Liu X, Si W, Garau C, Jüngling K, Pape H-C, Schulz S, Reinscheid RK. 2017. Neuropeptide S precursor knockout mice display memory and arousal deficits. The European journal of neuroscience, 46 (1): 1689–1700. DOI 10.1111/ejn.13613.

Llano Lopez L, Hauser J, Feldon J, Gargiulo PA, Yee BK. 2010. Evaluating spatial memory function in mice: A within-subjects comparison between the water maze test and its adaptation to dry land. Behavioural Brain Research, 209 (1): 85–92. DOI 10.1016/j.bbr.2010.01.020.

Logue SF, Paylor R, Wehner JM. 1997. Hippocampal lesions cause learning deficits in inbred mice in the Morris water maze and conditioned-fear task. Behavioral neuroscience, 111 (1): 104–113. DOI 10.1037//0735-7044.111.1.104.

Lopes DM, Denk F, Chisholm KI, Suddason T, Durrieux C, Thakur M, Gentry C, McMahon SB. 2017. Peripheral inflammatory pain sensitisation is independent of mast cell activation in male mice. Pain, 158 (7): 1314–1322. DOI 10.1097/j.pain.0000000000000917.

López-Cano M, Fernández-Dueñas V, Llebaria A, Ciruela F. 2017. Formalin Murine Model of Pain. Bio-protocol, 7 (23): e2628. DOI 10.21769/BioProtoc.2628.

Lubow RE. 1973. Latent inhibition. Psychological bulletin, 79 (6): 398–407. DOI 10.1037/h0034425.

Luca A de. 2008. Use of grip strength meter to assess the limb strength of mdx mice. http://www.treat-nmd.eu/downloads/file/sops/dmd/MDX/DMD_M.2.2.001.pdf.

Luijten IHN, Feldmann HM, Essen G von, Cannon B, Nedergaard J. 2019. In the absence of UCP1-mediated diet-induced thermogenesis, obesity is augmented even in the obesity-resistant 129S mouse strain. American journal of physiology. Endocrinology and metabolism, 316 (5): E729-E740. DOI 10.1152/ajpendo.00020.2019.

Luong TN, Carlisle HJ, Southwell A, Patterson PH. 2011. Assessment of motor balance and coordination in mice using the balance beam. Journal of visualized experiments : JoVE, (49): 2376. DOI 10.3791/2376.

Lustig C, Kozak R, Sarter M, Young JW, Robbins TW. 2013. CNTRICS final animal model task selection: control of attention. Neuroscience & Biobehavioral Reviews, 37 (9 Pt B): 2099–2110. DOI 10.1016/j.neubiorev.2012.05.009.

Luttinger D. 1985. Determination of antinociceptive efficacy of drugs in mice using different water temperatures in a tail-immersion test. Journal of pharmacological methods, 13 (4): 351–357. DOI 10.1016/0160-5402(85)90017-8.

Mandairon N, Sacquet J, Garcia S, Ravel N, Jourdan F, Didier A. 2006. Neurogenic correlates of an olfactory discrimination task in the adult olfactory bulb. The European journal of neuroscience, 24 (12): 3578–3588. DOI 10.1111/j.1460-9568.2006.05235.x.

Marianno P, Abrahao KP, Camarini R. 2017. Environmental Enrichment Blunts Ethanol Consumption after Restraint Stress in C57BL/6 Mice. PloS one, 12 (1): e0170317. DOI 10.1371/journal.pone.0170317.

Marsch R, Foeller E, Rammes G, Bunck M, Kössl M, Holsboer F, Zieglgänsberger W, Landgraf R, Lutz B, Wotjak CT. 2007. Reduced anxiety, conditioned fear, and hippocampal long-term potentiation in transient receptor potential vanilloid type 1 receptor-deficient mice. The Journal of neuroscience : the official journal of the Society for Neuroscience, 27 (4): 832–839. DOI 10.1523/JNEUROSCI.3303-06.2007.

Martin AL, Brown RE. 2010. The lonely mouse: verification of a separation-induced model of depression in female mice. Behavioural Brain Research, 207 (1): 196–207. DOI 10.1016/j.bbr.2009.10.006.

Martin L, Iceberg E. 2015. Quantifying Social Motivation in Mice Using Operant Conditioning. Journal of visualized experiments : JoVE, (102): e53009. DOI 10.3791/53009.

Martinez M, Salvador A, Simón VM. 1994. Behavioral changes over several successful agonistic encounters between male mice: Effects of type of "standard opponent". Aggressive Behavior, 20 (6): 441–451. DOI 10.1002/1098-2337(1994)20:6<441::AID-AB2480200604>3.0.CO;2-W.

Masocha W, Kombian SB, Edafiogho IO. 2016. Evaluation of the antinociceptive activities of enaminone compounds on the formalin and hot plate tests in mice. Scientific Reports, 6: 21582. DOI 10.1038/srep21582.

Matsumoto K, Fujiwara H, Araki R, Yabe T. 2019. Post-weaning social isolation of mice: A putative animal model of developmental disorders. Journal of Pharmacological Sciences, 141 (3): 111–118. DOI 10.1016/j.jphs.2019.10.002.

Matsumoto YK, Okanoya K. 2016. Phase-Specific Vocalizations of Male Mice at the Initial Encounter during the Courtship Sequence. PloS one, 11 (2): e0147102. DOI 10.1371/journal.pone.0147102.

McLeod SA. 2018. https://www.simplypsychology.org/operant-conditioning.html (accessed Jul 8, 2020).

Mehla J, Deibel SH, Faraji J, Saito T, Saido TC, Mohajerani MH, McDonald RJ. 2019. Looking beyond the standard version of the Morris water task in the assessment of mouse models of cognitive deficits. Hippocampus, 29 (1): 3–14. DOI 10.1002/hipo.22999.

Mei J, Riedel N, Grittner U, Endres M, Banneke S, Emmrich JV. 2018. Body temperature measurement in mice during acute illness: implantable temperature transponder versus surface infrared thermometry. Scientific Reports, 8 (1): 3526. DOI 10.1038/s41598-018-22020-6.

Menéndez L, Lastra A, Hidalgo A, Baamonde A. 2002. Unilateral hot plate test: a simple and sensitive method for detecting central and peripheral hyperalgesia in mice. Journal of neuroscience methods, 113 (1): 91–97.

Meyer CW, Ootsuka Y, Romanovsky AA. 2017. Body Temperature Measurements for Metabolic Phenotyping in Mice. Frontiers in physiology, 8: 520.
DOI 10.3389/fphys.2017.00520.

Miao-Kun Sun, Hrsg. 2007. New Research in Cognitive Sciences. UK edition. Nova Science Publishers Inc.

Milanovic S, Radulovic J, Laban O, Stiedl O, Henn F, Spiess J. 1998. Production of the Fos protein after contextual fear conditioning of C57BL/6N mice. Brain research, 784 (1-2): 37–47. DOI 10.1016/s0006-8993(97)01266-3.

Miller AL, Leach MC. 2014. Using the mouse grimace scale to assess pain associated with routine ear notching and the effect of analgesia in laboratory mice. Laboratory animals, 49 (2): 117–120. DOI 10.1177/0023677214559084.

Miller AL, Leach MC. 2015. The Mouse Grimace Scale: A Clinically Useful Tool? PloS one, 10 (9): e0136000. DOI 10.1371/journal.pone.0136000.

Minett MS, Eijkelkamp N, Wood JN. 2014. Significant determinants of mouse pain behaviour. PloS one, 9 (8): e104458. DOI 10.1371/journal.pone.0104458.

Mogi K, Takakuda A, Tsukamoto C, Ooyama R, Okabe S, Koshida N, Nagasawa M, Kikusui T. 2017. Mutual mother-infant recognition in mice: The role of pup ultrasonic vocalizations. Behavioural Brain Research, 325: 138–146. DOI 10.1016/j.bbr.2016.08.044.

Montilla-Garcia A, Tejada MA, Perazzoli G, Entrena JM, Portillo-Salido E, Fernandez-Segura E, Canizares FJ, Cobos EJ. 2017. Grip strength in mice with joint inflammation: A

rheumatology function test sensitive to pain and analgesia. Neuropharmacology, 125: 231–242. DOI 10.1016/j.neuropharm.2017.07.029.

Moy SS, Nadler JJ, Perez A, Barbaro RP, Johns JM, Magnuson TR, Piven J, Crawley JN. 2004. Sociability and preference for social novelty in five inbred strains: an approach to assess autistic-like behavior in mice. Genes, Brain and Behavior, 3 (5): 287–302. DOI 10.1111/j.1601-1848.2004.00076.x.

Muller CR, Américo ALV, Fiorino P, Evangelista FS. 2018. Aerobic exercise training prevents kidney lipid deposition in mice fed a cafeteria diet. Life sciences, 211: 140–146. DOI 10.1016/j.lfs.2018.09.017.

Nakamoto K, Nishinaka T, Mankura M, Fujita-Hamabe W, Tokuyama S. 2010. Antinociceptive effects of docosahexaenoic acid against various pain stimuli in mice. Biological & pharmaceutical bulletin, 33 (6): 1070–1072. DOI 10.1248/bpb.33.1070.

National Research Council. 2009. Recognition and Alleviation of Pain in Laboratory Animals. Washington, DC: The National Academies Press.

Nevins ME, Nash SA, Beardsley PM. 1993. Quantitative grip strength assessment as a means of evaluating muscle relaxation in mice. Psychopharmacology, 110 (1-2): 92–96. DOI 10.1007/bf02246955.

Newman EL, Covington HE, Suh J, Bicakci MB, Ressler KJ, DeBold JF, Miczek KA. 2019. Fighting Females: Neural and Behavioral Consequences of Social Defeat Stress in Female Mice. Biological psychiatry, 86 (9): 657–668. DOI 10.1016/j.biopsych.2019.05.005.

Nikoui V, Ostadhadi S, Imran Khan M, Allahverdi A, Bakhtiarian* A. 2016. Evaluation of the analgesic effect of Thymus kotschyanus hydroalcoholic extract in male mice by formalin and tail flick tests. Journal of Basic and Clinical Pathophysiology, 4 (1): 1–6. DOI 10.22070/jbcp.2016.277.

Nuzzo D, Amato A, Picone P, Terzo S, Galizzi G, Bonina FP, Mulè F, Di Carlo M. 2018. A Natural Dietary Supplement with a Combination of Nutrients Prevents Neurodegeneration Induced by a High Fat Diet in Mice. Nutrients, 10 (9). DOI 10.3390/nu10091130.

O. H. Maghsoudi, A. V. Tabrizi, B. Robertson, P. Shamble, A. Spence. 2015. A novel automatic method to track the body and paws of running mice in high speed video. In: . 2015 IEEE Signal Processing in Medicine and Biology Symposium (SPMB). : 1–2.

O'Brien DE, Brenner DS, Gutmann DH, Gereau RW, 4th. 2013. Assessment of pain and itch behavior in a mouse model of neurofibromatosis type 1. The journal of pain : official journal of the American Pain Society, 14 (6): 628–637. DOI 10.1016/j.jpain.2013.01.770.

Okabe S, Nagasawa M, Kihara T, Kato M, Harada T, Koshida N, Mogi K, Kikusui T. 2013. Pup odor and ultrasonic vocalizations synergistically stimulate maternal attention in mice. Behavioral neuroscience, 127 (3): 432–438. DOI 10.1037/a0032395.

O'Leary TP, Brown RE. 2009. Visuo-spatial learning and memory deficits on the Barnes maze in the 16-month-old APPswe/PS1dE9 mouse model of Alzheimer's disease. Behavioural Brain Research, 201 (1): 120–127. DOI 10.1016/j.bbr.2009.01.039.

O'Leary TP, Brown RE. 2012. The effects of apparatus design and test procedure on learning and memory performance of C57BL/6J mice on the Barnes maze. Journal of neuroscience methods, 203 (2): 315–324. DOI 10.1016/j.jneumeth.2011.09.027.

Oliván S, Calvo AC, Rando A, Muñoz MJ, Zaragoza P, Osta R. Comparative study of behavioural tests in the SOD1G93A mouse model of amyotrophic lateral sclerosis. Experimental animals, 64 (2): 147–153. DOI 10.1538/expanim.14-0077.

Olton DS, Samuelson RJ. 1976. Remembrance of places passed: Spatial memory in rats. Journal of Experimental Psychology: Animal Behavior Processes, 2 (2): 97–116. DOI 10.1037/0097-7403.2.2.97.

Otsuka S. 2017. Predator Odor-induced Freezing Test for Mice. Bio-protocol, 7 (17): e2534. DOI 10.21769/BioProtoc.2534.

Padgett DA, Marucha PT, Sheridan JF. 1998. Restraint Stress Slows Cutaneous Wound Healing in Mice. Brain, Behavior, and Immunity, 12 (1): 64–73. DOI 10.1006/brbi.1997.0512.

Paiva VN de, Lima SNP, Fernandes MM, Soncini R, Andrade CAF, Giusti-Paiva A. 2010. Prostaglandins mediate depressive-like behaviour induced by endotoxin in mice. Behavioural Brain Research, 215 (1): 146–151. DOI 10.1016/j.bbr.2010.07.015.

Parvathy SS, Masocha W. 2013. Gait analysis of C57BL/6 mice with complete Freund's adjuvant-induced arthritis using the CatWalk system. BMC musculoskeletal disorders, 14: 14. DOI 10.1186/1471-2474-14-14.

Patchev VK, Patchev AV. 2006. Experimental models of stress. Dialogues in clinical neuroscience, 8 (4): 417–432. https://pubmed.ncbi.nlm.nih.gov/17290800.

Patil SS, Sunyer B, Hoger H, Lubec G. 2009. Evaluation of spatial memory of C57BL/6J and CD1 mice in the Barnes maze, the Multiple T-maze and in the Morris water maze. Behavioural Brain Research, 198 (1): 58–68. DOI 10.1016/j.bbr.2008.10.029.

Pawlyk AC, Morrison AR, Ross RJ, Brennan FX. 2008. Stress-induced changes in sleep in rodents: Models and mechanisms. Neuroscience & Biobehavioral Reviews, 32 (1): 99–117. DOI 10.1016/j.neubiorev.2007.06.001.

Pearson BL, Defensor EB, Blanchard DC, Blanchard RJ. 2010. C57BL/6J mice fail to exhibit preference for social novelty in the three-chamber apparatus. Behavioural Brain Research, 213 (2): 189–194. DOI 10.1016/j.bbr.2010.04.054.

Pereira C, Kunczik J, Bleich A, Haeger C, Kiessling F, Thum T, Tolba R, Lindauer U, Treue S, Czaplik M. 2019. Perspective review of optical imaging in welfare assessment in animal-based research. Journal of biomedical optics, 24 (7): 1–11. DOI 10.1117/1.JBO.24.7.070601.

Pereira CB, Kunczik J, Zieglowski L, Tolba R, Abdelrahman A, Zechner D, Vollmar B, Janssen H, Thum T, Czaplik M. 2018. Remote Welfare Monitoring of Rodents Using Thermal Imaging. Sensors (Basel, Switzerland), 18 (11). DOI 10.3390/s18113653.

Perepelkina OV, Tarassova AY, Surina NM, Lilp IG, Golibrodo VA, Poletaeva II. 2017. Intermale aggression in mice, selected for the cognitive trait. Doklady biological sciences : proceedings of the Academy of Sciences of the USSR, Biological sciences sections, 475 (1): 151–153. DOI 10.1134/S001249661704010X.

Pierce RC, Kalivas PW. 2007. Locomotor behavior. Current Protocols in Neuroscience, Chapter 8: Unit 8.1. DOI 10.1002/0471142301.ns0801s40.

Pietropaolo S, Singer P, Feldon J, Yee BK. 2008. The postweaning social isolation in C57BL/6 mice: preferential vulnerability in the male sex. Psychopharmacology, 197 (4): 613–628. DOI 10.1007/s00213-008-1081-3.

Pitts MW. 2018. Barnes Maze Procedure for Spatial Learning and Memory in Mice. Bio-protocol, 8 (5): e2744. DOI 10.21769/bioprotoc.2744.

Pollak DD, Monje FJ, Lubec G. 2010. The learned safety paradigm as a mouse model for neuropsychiatric research. Nature protocols, 5: 954 EP -. DOI 10.1038/nprot.2010.64.

Pomerantz SM, Nunez AA, Jay Bean N. 1983. Female behavior is affected by male ultrasonic vocalizations in house mice. Physiology & Behavior, 31 (1): 91–96. DOI 10.1016/0031-9384(83)90101-4.

Ponnusamy R, Nissim HA, Barad M. 2005. Systemic blockade of D2-like dopamine receptors facilitates extinction of conditioned fear in mice. Learning & memory (Cold Spring Harbor, N.Y.), 12 (4): 399–406. DOI 10.1101/lm.96605.

Porsolt RD, Le Pichon M, Jalfre M. 1977. Depression: a new animal model sensitive to antidepressant treatments. Nature, 266 (5604): 730–732. DOI 10.1038/266730a0.

Prado VF, Martins-Silva C, Castro BM de, Lima RF, Barros DM, Amaral E, Ramsey AJ, Sotnikova TD, Ramirez MR, Kim H-G, Rossato JI, Koenen J, Quan H, Cota VR, Moraes MFD, Gomez MV, Guatimosim C, Wetsel WC, Kushmerick C, Pereira GS, Gainetdinov RR, Izquierdo I, Caron MG, Prado MAM. 2006. Mice deficient for the vesicular acetylcholine transporter are myasthenic and have deficits in object and social recognition. Neuron, 51 (5): 601–612. DOI 10.1016/j.neuron.2006.08.005.

Prusky GT, West PWR, Douglas RM. 2000. Behavioral assessment of visual acuity in mice and rats. Vision Research, 40 (16): 2201–2209. DOI 10.1016/S0042-6989(00)00081-X.

Prut L, Belzung C. 2003. The open field as a paradigm to measure the effects of drugs on anxiety-like behaviors: a review. European journal of pharmacology, 463 (1): 3–33. DOI 10.1016/S0014-2999(03)01272-X.

Qu C, Xiong Y, Mahmood A, Kaplan DL, Goussev A, Ning R, Chopp M. 2009. Treatment of traumatic brain injury in mice with bone marrow stromal cell-impregnated collagen scaffolds. Journal of neurosurgery, 111 (4): 658–665. DOI 10.3171/2009.4.JNS081681.

Radulovic J, Kammermeier J, Spiess J. 1998. Relationship between fos production and classical fear conditioning: effects of novelty, latent inhibition, and unconditioned stimulus preexposure. The Journal of neuroscience : the official journal of the Society for Neuroscience, 18 (18): 7452–7461. DOI 10.1523/JNEUROSCI.18-18-07452.1998.

Ramabadran K, Bansinath M, Turndorf H, Puig MM. 1989. Tail immersion test for the evaluation of a nociceptive reaction in mice: Methodological considerations. Journal of pharmacological methods, 21 (1): 21–31. DOI 10.1016/0160-5402(89)90019-3.

Rammal H, Bouayed J, Soulimani R. 2010. A direct relationship between aggressive behavior in the resident/intruder test and cell oxidative status in adult male mice. European journal of pharmacology, 627 (1): 173–176. DOI 10.1016/j.ejphar.2009.11.001.

Rasmussen A, Ijpelaar ACHG, Zeeuw CI de, Boele H-J. 2018. Caffeine has no effect on eyeblink conditioning in mice. Behavioural Brain Research, 337: 252–255. DOI 10.1016/j.bbr.2017.09.013.

Robbins TW. 2002. The 5-choice serial reaction time task: behavioural pharmacology and functional neurochemistry. Psychopharmacology, 163 (3-4): 362–380. DOI 10.1007/s00213-002-1154-7.

Roemers P, Mazzola PN, Deyn PP de, Bossers WJ, van Heuvelen, M J G, van der Zee, E A. 2018. Burrowing as a novel voluntary strength training method for mice: A comparison of various voluntary strength or resistance exercise methods. Journal of neuroscience methods, 300: 112–126. DOI 10.1016/j.jneumeth.2017.05.027.

Rogers DC, Fisher EM, Brown SD, Peters J, Hunter AJ, Martin JE. 1997. Behavioral and functional analysis of mouse phenotype: SHIRPA, a proposed protocol for comprehensive phenotype assessment. Mammalian genome : official journal of the International Mammalian Genome Society, 8 (10): 711–713. DOI 10.1007/s003359900551.

Rogoyski A. 1966. Short lasting food deprivation in pregnant mice as a factor responsible for cleft palate in the outcomes. Ginek. Pol., (37): 1049–1052.

Roome RB, Vanderluit JL. 2015. Paw-dragging: a novel, sensitive analysis of the mouse cylinder test. Journal of visualized experiments : JoVE, (98): e52701. DOI 10.3791/52701.

Rosenzweig S, Blaustein FM. 1970. Cleft palate in A/J mice resulting from restraint and deprivation of food and water. Teratology, 3 (1): 47–52. DOI 10.1002/tera.1420030110.

Rosland JH, Tjølsen A, Maehle B, Hole K. 1990. The formalin test in mice: effect of formalin concentration. Pain, 42 (2): 235–242. DOI 10.1016/0304-3959(90)91167-h.

Runner, M. N., Miller J. R. 1956. Congenital deformity in the mouse as a consequence f fasting. The Anatomical Record, (124): 437–438.

Sabbath EL, Gutierrez L-A, Okechukwu CA, Singh-Manoux A, Amieva H, Goldberg M, Zins M, Berr C. 2014. Time may not fully attenuate solvent-associated cognitive deficits in highly exposed workers. Neurology, 82 (19): 1716–1723. DOI 10.1212/WNL.0000000000000413.

Saegusa Y, Tabata H. 2003. Usefulness of infrared thermometry in determining body temperature in mice. The Journal of veterinary medical science, 65 (12): 1365–1367. DOI 10.1292/jvms.65.1365.

Saha S, Guria T, Singha T, Maity TK. 2013. Evaluation of Analgesic and Anti-Inflammatory Activity of Chloroform and Methanol Extracts of Centella asiatica Linn. ISRN pharmacology, 2013: 789613. DOI 10.1155/2013/789613.

Sakata A, Mogi M, Iwanami J, Tsukuda K, Min L-J, Fujita T, Iwai M, Ito M, Horiuchi M. 2009. Sex-different effect of angiotensin II type 2 receptor on ischemic brain injury and cognitive function. Brain research, 1300: 14–23. DOI 10.1016/j.brainres.2009.08.068.

Sampey BP, Vanhoose AM, Winfield HM, Freemerman AJ, Muehlbauer MJ, Fueger PT, Newgard CB, Makowski L. 2011. Cafeteria diet is a robust model of human metabolic syndrome with liver and adipose inflammation: comparison to high-fat diet. Obesity (Silver Spring, Md.), 19 (6): 1109–1117. DOI 10.1038/oby.2011.18.

Samuels B.A. HR. 2011. Novelty-Supressed Feeding in the Mouse. Mood and Anxiety Related Phenotypes in Mice, Neuromethods, (63). DOI 10.1007/978-1-61779-313-4_7.

Santos-Nogueira E, Redondo Castro E, Mancuso R, Navarro X. 2012. Randall-Selitto test: a new approach for the detection of neuropathic pain after spinal cord injury. Journal of neurotrauma, 29 (5): 898–904. DOI 10.1089/neu.2010.1700.

Schefer V, Talan MI. 1996. Oxygen consumption in adult and aged C57BL/6J mice during acute treadmill exercise of different intensity. Experimental Gerontology, 31 (3): 387–392. DOI 10.1016/0531-5565(95)02032-2.

Scholz J, Niibori Y, W Frankland P, P Lerch J. 2015. Rotarod training in mice is associated with changes in brain structure observable with multimodal MRI. NeuroImage, 107: 182–189. DOI 10.1016/j.neuroimage.2014.12.003.

Sclafani A, Glendinning JI. 2003. Flavor preferences conditioned in C57BL/6 mice by intragastric carbohydrate self-infusion. Physiology & Behavior, 79 (4): 783–788. DOI 10.1016/S0031-9384(03)00174-4.

Serchov T, van Calker D, Biber K. 2016. Light/Dark Transition Test to Assess Anxiety-like Behavior in Mice. Bio-protocol, 6 (19). DOI 10.21769/BioProtoc.1957.

Shelton SB, Pettigrew DB, Hermann AD, Zhou W, Sullivan PM, Crutcher KA, Strauss KI. 2008. A simple, efficient tool for assessment of mice after unilateral cortex injury. Journal of neuroscience methods, 168 (2): 431–442. DOI 10.1016/j.jneumeth.2007.11.003.

Shepherd JK, Grewal SS, Fletcher A, Bill DJ, Dourish CT. 1994. Behavioural and pharmacological characterisation of the elevated "zero-maze" as an animal model of anxiety. Psychopharmacology, 116 (1): 56–64. DOI 10.1007/BF02244871.

Sherwin CM. 1998. Voluntary wheel running: a review and novel interpretation. Animal Behaviour, 56 (1): 11–27. DOI 10.1006/anbe.1998.0836.

Sherwin CM, Christiansen SB, Duncan IJ, Erhard HW, Lay DC, Mench JA, O'Connor CE, Petherick JC. 2003. Guidelines for the ethical use of animals in applied ethology studies. Applied Animal Behaviour Science, 81 (3): 291–305. DOI 10.1016/S0168-1591(02)00288-5.

Sheth KA, Iyer CC, Wier CG, Crum AE, Bratasz A, Kolb SJ, Clark BC, Burghes AHM, Arnold WD. 2018. Muscle strength and size are associated with motor unit connectivity in aged mice. Neurobiology of aging, 67: 128–136. DOI 10.1016/j.neurobiolaging.2018.03.016.

Shiotsuki H, Yoshimi K, Shimo Y, Funayama M, Takamatsu Y, Ikeda K, Takahashi R, Kitazawa S, Hattori N. 2010. A rotarod test for evaluation of motor skill learning. Journal of neuroscience methods, 189 (2): 180–185. DOI 10.1016/j.jneumeth.2010.03.026.

Shukitt–Hale B, Smith DE, Meydani M, Joseph JA. 1999. The effects of dietary antioxidants on psychomotor performance in aged mice. Experimental Gerontology, 34 (6): 797–808. DOI 10.1016/S0531-5565(99)00039-X.

Siopi E, Denizet M, Gabellec M-M, Chaumont F de, Olivo-Marin J-C, Guilloux J-P, Lledo P-M, Lazarini F. 2016. Anxiety- and Depression-Like States Lead to Pronounced Olfactory Deficits and Impaired Adult Neurogenesis in Mice. The Journal of neuroscience : the official journal of the Society for Neuroscience, 36 (2): 518–531. DOI 10.1523/JNEUROSCI.2817-15.2016.

Sotocinal SG, Sorge RE, Zaloum A, Tuttle AH, Martin LJ, Wieskopf JS, Mapplebeck JCS, Wei P, Zhan S, Zhang S, McDougall JJ, King OD, Mogil JS. 2011. The Rat Grimace Scale: a partially automated method for quantifying pain in the laboratory rat via facial expressions. Molecular pain, 7: 55. DOI 10.1186/1744-8069-7-55.

Sousa N, Almeida OFX, Wotjak CT. 2006. A hitchhiker's guide to behavioral analysis in laboratory rodents. Genes, brain, and behavior, 5 Suppl 2: 5–24. DOI 10.1111/j.1601-183X.2006.00228.x.

Spangenberg EMF, Wichman A. 2018. Methods for Investigating the Motivation of Mice to Explore and Access Food Rewards. Journal of the American Association for Laboratory Animal Science : JAALAS, 57 (3): 244–252. DOI 10.30802/AALAS-JAALAS-17-000080.

Stanford Medicine Behavioral and Functional Neuroscience Laboratory. 2020. https://med.stanford.edu/sbfnl/services/bm/sm/BasketTest.html (accessed Jun 29, 2020).

Stanley JL, Lincoln RJ, Brown TA, McDonald LM, Dawson GR, Reynolds DS. 2005. The mouse beam walking assay offers improved sensitivity over the mouse rotarod in deter-

mining motor coordination deficits induced by benzodiazepines. Journal of psychopharmacology (Oxford, England), 19 (3): 221–227. DOI 10.1177/0269881105051524.

Starkey ML, Barritt AW, Yip PK, Davies M, Hamers FPT, McMahon SB, Bradbury EJ. 2005. Assessing behavioural function following a pyramidotomy lesion of the corticospinal tract in adult mice. Experimental Neurology, 195 (2): 524–539. DOI 10.1016/j.expneurol.2005.06.017.

Sterniczuk R, Antle MC, Laferla FM, Dyck RH. 2010. Characterization of the 3xTg-AD mouse model of Alzheimer's disease: part 2. Behavioral and cognitive changes. Brain research, 1348: 149–155. DOI 10.1016/j.brainres.2010.06.011.

Steru L, Chermat R, Thierry B, Simon P. 1985. The tail suspension test: a new method for screening antidepressants in mice. Psychopharmacology, 85 (3): 367–370. DOI 10.1007/BF00428203.

Stowers JR, Hofbauer M, Bastien R, Griessner J, Higgins P, Farooqui S, Fischer RM, Nowikovsky K, Haubensak W, Couzin ID, Tessmar-Raible K, Straw AD. 2017. Virtual reality for freely moving animals. Nature methods, 14 (10): 995–1002. DOI 10.1038/nmeth.4399.

Sulakhiya K, Patel VK, Saxena R, Dashore J, Srivastava AK, Rathore M. 2016. Effect of Beta vulgaris Linn. Leaves Extract on Anxiety- and Depressive-like Behavior and Oxidative Stress in Mice after Acute Restraint Stress. Pharmacognosy research, 8 (1): 1–7. DOI 10.4103/0974-8490.171100.

Tachikawa KS, Yoshihara Y, Kuroda KO. 2013. Behavioral transition from attack to parenting in male mice: a crucial role of the vomeronasal system. The Journal of neuroscience : the official journal of the Society for Neuroscience, 33 (12): 5120–5126. DOI 10.1523/JNEUROSCI.2364-12.2013.

Takahashi A, Yap JJ, Bohager DZ, Faccidomo S, Clayton T, Cook JM, Miczek KA. 2009. Glutamatergic and GABAergic modulations of ultrasonic vocalizations during maternal separation distress in mouse pups. Psychopharmacology, 204 (1): 61–71. DOI 10.1007/s00213-008-1437-8.

Takeda H, Tsuji M, Matsumiya T. 1998. Changes in head-dipping behavior in the hole-board test reflect the anxiogenic and/or anxiolytic state in mice. European journal of pharmacology, 350 (1): 21–29.

Tang X, Orchard SM, Sanford LD. 2002. Home cage activity and behavioral performance in inbred and hybrid mice. Behavioural Brain Research, 136 (2): 555–569. DOI 10.1016/s0166-4328(02)00228-0.

Tang X, Sanford LD. 2002. Telemetric recording of sleep and home cage activity in mice. Sleep, 25 (6): 691–699.

Teegarden S. 2012. Behavioral Phenotyping in Rats and Mice. Materials and Methods, 2. DOI 10.13070/mm.en.2.122.

Thomas D. Tyler, Richard E. Tessel. 1979. A new device for the simultaneous measurement of locomotor and stereotypic frequency in mice. Psychopharmacology, 64 (3): 285. DOI 10.1007/BF00427511.

Tillerson JL, Caudle WM, Parent JM, Gong C, Schallert T, Miller GW. 2006. Olfactory discrimination deficits in mice lacking the dopamine transporter or the D2 dopamine receptor. Behavioural Brain Research, 172 (1): 97–105. DOI 10.1016/j.bbr.2006.04.025.

Tjolsen A, Rosland JH, Berge OG, Hole K. 1991. The increasing-temperature hot-plate test: an improved test of nociception in mice and rats. Journal of pharmacological methods, 25 (3): 241–250.

Trueman RC, Brooks SP, Dunnett SB. 2005. Implicit learning in a serial choice visual discrimination task in the operant 9-hole box by intact and striatal lesioned mice. Behavioural Brain Research, 159 (2): 313–322. DOI 10.1016/j.bbr.2004.11.021.

Umezu T. 1999. Effects of psychoactive drugs in the Vogel conflict test in mice. Japanese journal of pharmacology, 80 (2): 111–118.

Valsamis B, Schmid S. 2011. Habituation and prepulse inhibition of acoustic startle in rodents. Journal of visualized experiments : JoVE, (55): e3446. DOI 10.3791/3446.

Valzelli L. 1973. The “isolation syndrome” in mice. Psychopharmacologia, 31 (4): 305–320. DOI 10.1007/BF00421275.

Valzelli L, Bernasconi S, Gomba P. 1974. Effect of isolation on some behavioral characteristics in three strains of mice. Biological psychiatry, 9 (3): 329–334.

van de Weerd HA, van den Broek FAR, Beynen AC. 1992. Removal of vibrissae in male mice does not influence social dominance. Behavioural Processes, 27 (3): 205–208. DOI 10.1016/0376-6357(92)90177-F.

van de Weerd HA, van Loo PL, van Zutphen LF, Koolhaas JM, Baumans V. 1997. Preferences for nesting material as environmental enrichment for laboratory mice. Laboratory animals, 31 (2): 133–143. DOI 10.1258/002367797780600152.

van Loo PLP, Kuin N, Sommer R, Avsaroglu H, Pham T, Baumans V. 2007. Impact of 'living apart together' on postoperative recovery of mice compared with social and individual housing. Laboratory animals, 41 (4): 441–455. DOI 10.1258/002367707782314328.

van Putten M. 2011. The use of hanging wire tests to monitor muscle strength and condition over time. http://www.treat-nmd.eu/downloads/file/sops/dmd/MDX/DMD_M.2.1.004.pdf (accessed June 29th, 2016).

Vareniuk I, Pavlov IA, Obrosova IG. 2008. Inducible nitric oxide synthase gene deficiency counteracts multiple manifestations of peripheral neuropathy in a streptozotocin-induced mouse model of diabetes. Diabetologia, 51 (11): 2126–2133. DOI 10.1007/s00125-008-1136-3.

Wahlsten D. 2011. Mouse behavioral testing. How to use mice in behavioral neuroscience. London: Elsevier, 288.

Walf AA, Frye CA. 2007. The use of the elevated plus maze as an assay of anxiety-related behavior in rodents. Nature protocols, 2 (2): 322–328. DOI 10.1038/nprot.2007.44.

Walker DL, Toufexis DJ, Davis M. 2003. Role of the bed nucleus of the stria terminalis versus the amygdala in fear, stress, and anxiety. European journal of pharmacology, 463 (1): 199–216. DOI 10.1016/S0014-2999(03)01282-2.

Watt LJ. 1931. OVULATION, OESTRUS AND COPULATION WITH CONSEQUENT DYSTOCIA DURING PREGNANCY, IN THE MOUSE. Science (New York, N.Y.), 73 (1881): 75–76. DOI 10.1126/science.73.1881.75.

Weidenauer A, Sauerzopf U, Bartova L, Bauer M, Praschak-Rieder N, Kasper S, Willeit M. 2018. Sensibilisierung auf Amphetamine: Bedeutung für die Schizophrenie- und Suchtforschung // Sensitization to amphetamines: Relevance for schizophrenia and addiction research. Journal für Neurologie, Neurochirurgie und Psychatrie, 19 (22): 66–71. https://www.kup.at/kup/pdf/14026.pdf.

Wen T, Ansonoff MA, Pintar JE. 2009. The tail pigmentation pattern of C57BL/6J mice affects nociception/pain quantification in the tail flick test. European journal of pain (London, England), 13 (6): 564–567. DOI 10.1016/j.ejpain.2008.07.012.

Willott JF. 2001. Handbook of mouse auditory research. From behavior to molecular biology. Boca Raton: CRC Press, 721.

Wilson SG, Mogil JS. 2001. Measuring pain in the (knockout) mouse: big challenges in a small mammal. Behavioural Brain Research, 125 (1): 65–73. DOI 10.1016/S0166-4328(01)00281-9.

Winslow JT. 2003. Mouse Social Recognition and Preference. Current Protocols in Neuroscience, 22 (1): 8.16.1-8.16.16. DOI 10.1002/0471142301.ns0816s22.

Woolfe G, Macdonald AD. 1944a. The Evaluation Of The Analgesic Action Of Pethidine Hydrochloride (Dermol). Journal of Pharmacology and Experimental Therapeutics, 80 (3): 300.

Woolfe G, Macdonald AD. 1944b. The Evaluation Of The Analgesic Action Of Pethidine Hydrochloride (Dermol). Journal of Pharmacology and Experimental Therapeutics, 80 (3): 300.

Wrenn CC. 2004. Social transmission of food preference in mice. Current Protocols in Neuroscience, Chapter 8: Unit 8.5G. DOI 10.1002/0471142301.ns0805gs28.

Xu L, Xiong X, Ouyang Y, Barreto G, Giffard R. 2011. Heat shock protein 72 (Hsp72) improves long term recovery after focal cerebral ischemia in mice. Neuroscience Letters, 488 (3): 279–282. DOI 10.1016/j.neulet.2010.11.047.

Yang M, Crawley JN. 2009. Simple behavioral assessment of mouse olfaction. Current Protocols in Neuroscience, Chapter 8: Unit-8.24. DOI 10.1002/0471142301.ns0824s48.

Yhnell E, Dunnett SB, Brooks SP. 2016. The utilisation of operant delayed matching and non-matching to position for probing cognitive flexibility and working memory in mouse models of Huntington's disease. Journal of neuroscience methods, 265: 72–80. DOI 10.1016/j.jneumeth.2015.08.022.

Zeeni N, Dagher-Hamalian C, Dimassi H, Faour WH. 2015. Cafeteria diet-fed mice is a pertinent model of obesity-induced organ damage: a potential role of inflammation. Inflammation research : official journal of the European Histamine Research Society … [et al.], 64 (7): 501–512. DOI 10.1007/s00011-015-0831-z.

Zhang L, Schallert T, Zhang ZG, Jiang Q, Arniego P, Li Q, Lu M, Chopp M. 2002. A test for detecting long-term sensorimotor dysfunction in the mouse after focal cerebral ischemia. Journal of neuroscience methods, 117 (2): 207–214.

Zhang M, Wang Y, Geng J, Zhou S, Xiao B. 2019. Mechanically Activated Piezo Channels Mediate Touch and Suppress Acute Mechanical Pain Response in Mice. Cell Reports, 26 (6): 1419-1431.e4. DOI 10.1016/j.celrep.2019.01.056.

Zulazmi NA, Gopalsamy B, Farouk AAO, Sulaiman MR, Bharatham BH, Perimal EK. 2015. Antiallodynic and antihyperalgesic effects of zerumbone on a mouse model of chronic constriction injury-induced neuropathic pain. Fitoterapia, 105: 215–221. DOI 10.1016/j.fitote.2015.07.011.

10. List of Figures and Tables

Fig. 1: Use of laboratory animals in Austria from 2007 – 2018

Tab. 1: Keywords used for literature research

Tab. 2: Tools used for literature research

Tab. 3: Composition of the used diets

Tab. 4: Overview of behavioural tests

11. Abbreviations

dB	Decibel
DMTP	Delayed-matching-to-position
DNMTP	Delayed-non-matching-to-position
FACS	Facial action coding system
g	Gramm; g-force
h	Hour
IMPC	International Mouse Phenotyping Consortium
kHz	Kilohertz
L	Liter
lx	Lux
mA	Milliamp
min	Minute
ml	Milliliter
rpm	Rotations per minute
s	Second
TVG2012	Tierversuchsgesetz 2012
W	Watt
µl	Microliter