Review

The complexities of behavioural assessment in neurodegenerative disorders: A focus on Alzheimer's disease

Sean Commins[a], Brian P. Kirby[b],*

[a] *Department of Psychology, Maynooth University, Co Kildare, Ireland*
[b] *School of Pharmacy, Royal College of Surgeons in Ireland, 123 St Stephen's Green, Dublin 2, Ireland*

ABSTRACT

Rodent models of human diseases that accurately and reproducibly capture their pathology are key tools in furthering our understanding of the mechanisms behind these diseases and in the development of novel treatment approaches. However, pre-clinical studies in rodents are often criticised for the relative lack of replication and success upon translation to humans. Animal models of neurodegenerative diseases (and other CNS conditions) are very complex, often with multifactorial inputs into their development and progression. This complexity is a significant challenge. In addition to this, there are often concerns raised about the conduct, analysis and interpretation of the model results. This review focuses on Alzheimer's disease as a representative neurodegenerative disorder and will examine disease model end-points, in particular, behavioural phenotyping which, while appearing relatively straightforward, has the potential to be poorly conducted and the results misconstrued. This review uses a sample of the literature to illustrate the breadth of techniques used in behavioural assessment of Alzheimer's disease models, highlight the complexity, illustrate some procedural, interpretational and translational issues and provide recommendations to improve conduct of pre-clinical testing with the hope that this may lead to more consistency and translational success.

1. Introduction

Neurodegenerative diseases encompass a range of disorders that affect our brain and may include Parkinson's, Alzheimer's, and Huntington's disease, as well as other disorders, such as prion disease, motor neurone disease and spinal muscular atrophy. In these conditions, cognitive and/or physical functionality progressively decreases in parallel with the neural degeneration, leading to severe disability and death. With an ageing population worldwide, the prevalence of neurodegenerative diseases is expected to rise over the coming years, representing a significant burden on health systems. According to the EU Joint Programme – Neurodegenerative Disease Research (JPND), dementias are responsible for the greatest burden of neurodegenerative diseases. For example, it is estimated that there are currently almost 47 million people worldwide living with dementia, the most common form of which is Alzheimer's disease (AD) [1]. Although the various neurodegenerative disorders have different clinical presentation and pathological changes, some of the underlying processes are common, such as protein deposition, mitochondrial dysfunction and excitotoxicity [2–4]. However, the process of neurodegeneration is complex and multifactorial and as a result, there are still no treatments to 'cure' AD or other neurodegenerative diseases. Despite this, there is significant ongoing research to try and better understand the pathological mechanisms so that new treatments and strategies can be developed.

Animal models form a critical experimental, pre-clinical step in order to further our knowledge of neurodegeneration, as well as, developing and testing treatments. The models aim to represent the particular disease in the animal, which then allows insights into the pathophysiology of the diseases. In addition to this, they are important in the assessment of efficacy of putative therapeutic agents before conducting human trials. This translation from animal models into humans has represented a significant challenge and there are far too many supposed treatments that have failed to be effective in humans. This failure is across all of the neurodegenerative conditions from Alzheimer's disease, where the much heralded anti-Aβ (beta-amyloid peptide) antibodies failed in clinical trials [5], to neuroprotection following cerebral ischaemia and the failure of NXY-059, a free radical-trapping agent, which showed a lack of efficacy in the Stroke Therapy Academic Industry Roundtable (STAIR) II trial [6]. One often mentioned reason for this translational failure is the poor design and reporting of pre-clinical studies (for review see [7]), leading to a lack of reproducibility and bias in reporting results [8,9]. However, as

---

mentioned above, it must be remembered that the complexity and multifactorial nature of neurodegenerative diseases is a significant challenge in translating results from the lab to the clinic. On the flip side, many clinical trials have not paid sufficient attention to the limitations of animal studies. This is evident in the stroke treatment trials whereby, for practical reasons, patients cannot be treated with the therapeutic agent inside the same time window (often 30–60 min post-ischaemia in animal studies) [10]. In order to increase the potential success of treatments translated from animals to humans, the conduct and reporting of animal studies needs to be improved. In addition, there are a number of key factors, applicable across all *in vivo* studies that should be considered. In this review, we will illustrate the general considerations before looking at AD in detail and some of the issues in the behavioural testing in AD models.

2. Key considerations in pre-clinical models of neurodegeneration

One of the key elements to improve the translation is that animal models must demonstrate validity. There are many applicable types of validity, including face and construct validity, whereby the behaviour of the animals mimics that of the condition (face validity) with a theoretical underpinning, such that the pathology, for example, is similar to that in patients (construct validity) [11]. However, more relevant in this review are internal and external validity. If an animal experiment has sufficient internal validity this means that, if investigating the effect of a novel treatment, the differences measured between groups of animals may be attributed to the novel treatment. Internal validity can be assured by careful study design and the minimisation of bias. The external validity refers to the extent to which a study reflects the disease in the human. Therefore, the closer a model is to the human condition, the better. A lack of external validity has been a factor in the failure of some treatments investigated for stroke whereby animals used in stroke studies are generally young and healthy but the patients who suffer a stroke often have co-morbidities, including hypertension or diabetes [12].

The choice of species, and indeed strain, is also a very important consideration. There is a large body of evidence showing the differences between, for example, rats and mice. While there is anatomical similarity between the brains of rats and mice, differences have been shown in neurotransmitter functions and localisation [13,14]. Furthermore, the rate of hippocampal neurogenesis has been shown to be greater in rats compared to mice and the newly generated cells are more likely to survive long-term and become involved in behaviour [15]. Similarly, while the vast majority of behavioural tests were originally designed for rats, they are now also used for mice on the assumption that the behaviour will be the same. However, species differences have been shown in, for example, water maze performance with mice performing more poorly than rats and also in tests of effort based decision making [16–18]. It is also significant to note that there are differences in behaviour between different strains of rodents. Much work has been done to look at different strains of rodents and the subtle differences in behaviour between them, which should be taken into account in modelling a disease in vivo [19,20]. For example, in the Morris water maze, some mouse strains are able to learn to find the platform more effectively than other strains [21].

Finally to add to the complexity in choosing the species and strain of animal for *in vivo* modelling, the gender is also important. There is evidence that the female brain responds to steroid hormones differently than the male brain which may impact stress responses and other behaviours [22]. Male mice have demonstrated better learning than female mice on both the working and reference memory components of a water-based radial arm maze task [23]. Furthermore, gender based differences in neurogenesis have been seen with steroid hormones modulating neurogenesis which may subsequently impact cognition and stress [24].

These animal-based factors have a direct impact on the behavioural

phenotype of a rodent. However, there are also environmental and human effects that can impact. For example, there is much evidence that differences in animal housing, both cages and their environment can greatly influence emotional responses and learning and memory in rodents [25,26]. Similarly, work by López-Aumatell and colleagues [27] demonstrated an experimenter effect on measures of anxiety in mice and other work has shown that handling of rodents can have an impact on responses to pharmacological treatment in the elevated plus maze test of anxiety [27,28]. In order to successfully translate results from the animal to the clinic, there must be consistency in approaches and results in animal experiments. One study that clearly illustrates the difficulty with environmental influences was carried out by Crabbe and colleagues [29]. They examined six different mouse behaviours across three different laboratories, while standardising all protocols, apparatus and animal environment yet still showed significant differences in test results between the labs.

What is clear from this is that there are a number of considerations that need to be taken in to account regarding *in vivo* modelling. Coupled with the complexity of neurodegeneration, the plethora of different models that can be chosen from and the specific phenotypic measures, the use of animals in modelling CNS disorders is something that has to be done carefully. Given this range and complexity, it is difficult for one review to deal comprehensively with neurodegeneration, or even behavioural testing, as a whole. Consequently, we have focussed on Alzheimer's disease and more specifically the behavioural assays that are utilised with AD models.

3. Alzheimer's disease

Alzheimer's disease (AD) is the most common form of dementia, accounting up to 60–70% of dementia case [30]. It is a devastating neurological disorder that affects an estimated 47 million people globally [31]. Worse than this, this figure is set to increase to 135.46 million by 2050. It is a global epidemic, not confined to Western populations; 63% of people with dementia live in low- and middle-income countries where support and social structures are often limited [31]. Furthermore, the disease is progressive, with the numbers of people living with dementia increasing with increased life-span: the numbers doubles with every 5.9 years increase in age, from 3.1/ 1000 person years at age 60–64, to 175.0/ 1000 person years at age 95 + [31]. Dementia and cognitive impairment are the leading cause of disability and dependence in older adults, with patients progressively unable to carry out routine but essential daily tasks, as well as, being unable to meet personal care needs [32]. The earliest symptoms of AD are impairments in memory, specifically episodic memory (a memory related to a particular event in one's life; when something took place, where did it occur, what happened and who was present) [30], followed by declarative (explicit memory for facts and events, such as knowing the capital of France) and non-declarative memory (implicit memory for motor habits and skills, such as riding a bicycle). In addition to memory impairments, other cognitive domains such as complex attention, executive functioning, language, perceptual-motor function and social cognition form the basis of diagnosis of dementia (or neurocognitive disorder (NCD) - Diagnostic and Statistical Manual of Mental Disorders (DSM-5) [33]).

At the neural level, AD is characterised by the presence of beta-amyloid peptide (Aβ) plaques that accumulate outside the cell and intracellular neurofibrillary tangles, which, in turn can lead to neural death. Beta-amyloid is a small peptide that is cleaved from the larger amyloid precursor protein (APP), which outside accumulates cells as a result of aberrant processing [34]. Intracellular neurofibrillary tangles are composed of hyperphosphorylated tau protein, a microtubule-associated protein (MAP), that under normal circumstances helps in the transport of organelles within the neuron as well as aiding in microtubular assembly [35]. However, with AD, tau accumulates in a hyperphosphorylated tangled form in the soma and neurites of the neuron

[30]. Neurofibrillary tangles are strongly associated with cognitive decline and tend to show a specific pattern of spread. Tangles first appear in the entorhinal cortex and hippocampal regions, before spreading to the amygdala and the anterodorsal thalamus. From these regions, tangles reach other cortical areas causing gross neurodegeneration [36].

Most animal models of AD have been developed either through genetic means or by the direct central administration of β-amyloid. In doing so, these models try to mimic the pathological hallmarks of AD and the cognitive decline. Intracerebral infusion of β-amyloid results in learning and memory deficits, as well as, many of the neuropathological hallmarks of Alzheimer's disease, such as nueroinflammation and microglial activation [37]. However, one limitation of the infusion models is that the level of β-amyloid in the rodent brain is much greater that would be seen in the brain of Alzheimer's disease patients. Genetically modified models (primarily mice) are excellent for establishing patho-histological changes that often map cognitive impairments, as well as, showing age-dependent development of neuropathology. For example, the 5XFAD and APP/PS1 models show amyloid plaques, synaptic degeneration and cellular death. In addition, these models develop cognitive deficits [30]. The APP/PS1 model show reference memory deficits at 6 months of age and persist after this; the model also shows spatial working memory impairments earlier, at 3 months of age [38,39]. Working memory deficits also seem to appear by 3 months in the 5XFAD model [40]. However, some mouse models have shown cognitive decline prior to amyloid plaques in the CNS [41,42]. For example, King and colleagues showed that the APP$_{SW}$ transgenic mouse exhibited early cognitive deficits, prior to amyloid plaque formation [43]. In addition, these aforementioned models lack tau pathology – a critical element in understanding AD [30]. Other models, such as the triple-transgenic AD (3xTgAD) mouse, do show region-specific Aβ plaques and tangles [44]; in addition, this model develops cognitive impairments, including associative learning deficits at 3–5 months of age, followed by working memory impairments at 6 months [39].

While there are different methods used to model AD, which we will not discuss here, the common denominator for these (and most of the neurodegenerative disorders) is behaviour. Behavioural tasks are routinely used to assess the extent or progression of neurodegeneration and the impact of potential therapeutic agents. In this review, we examine some of the more commonly used behavioural tests in AD and illustrate some differences between nominally identical tests that may confound the experiments and contribute to poor translation of results.

4. Behavioural assays commonly used with models of AD

To give a representative sample of the type of behavioural tasks that are commonly used with AD models, we conducted a brief search of the published literature. The intention was not to sample everything, thus producing a systematic review, but rather to obtain a representative sample of behavioural tasks used in AD models. The search strategy involved the use of two English language databases, namely PubMed and ScienceDirect, in February 2019. The search terms used were 'Alzheimer's disease', 'Behavioural tasks' and 'Animals' within the years from 2010 to 2018. Eighty-two articles were initially identified following the removal of repetitions. From this, a number of articles were not included as they were not directly relevant to our search criteria or the full text was unavailable; we were then left with 52 publications. We examined each of these papers and identified the species, sex, AD model, the behavioural task used and the functional domain associated with the task. From this, we found that the majority of papers examined cognitive functioning, as might be expected (44 papers), which were often combined with examination of other domains including general exploration (20 papers), motor functioning (15 papers) and anxiety (10 papers). Other behaviours such as sensory-motor gating, social interaction/discrimination, olfaction and visual discrimination were also examined but with much less frequency (Fig. 1a).

Within the *cognitive domain* the majority of researchers used the Morris water maze ([45] & Fig. 1a & b left) and this was used primarily as a test of spatial/reference memory and spatial working memory (e.g. [46] and Fig. 1a). The water maze task involves animals being placed into a pool of water for a number of trials. As rats/mice generally find water aversive, they try to escape it. Somewhere in the pool, hidden just below the surface, is a small platform, which the animals can find and escape on to. As the platform is invisible to the animal while swimming, it must use the various cues in the environment to find and recall the location of the platform [47]. The escape latency (i.e. the time to find the platform) and other parameters are used as a measure of spatial learning and memory. Depending on the training protocol, the water maze can also be used to examine spatial working memory and procedural memory [48]. For example, to test spatial working memory, many authors keep the platform location in the same position across trials within a day but change its position each day. Given that lesions of the hippocampus impair learning in this task [49,50], it measures important clinical features of AD, especially memory. Furthermore, the water maze is simple to use, is well validated and is a relatively cheap piece of equipment. Consequently, it is a very popular task in models of AD with most showing deficits (for example [51] with 3xTg-AD mice [52]; with APP/APLP2 double knockout (cDKO)). The radial arm maze and the radial water maze are also popular tasks to test spatial working memory [53]. The radial arm maze is an apparatus with several arms emanating from a central platform. At the end of each arm there is a food well which can contain a food reward. The animal must obtain the food from each arm, however the key feature is that on any given trial the animal must recall which arm it has previously visited and which arms still contain food. Deficits in these tasks have also been reported in various AD models including APP/PS-1 double knock-in [54] and PLB1Triple knock-in mice [55].

From our sample, the second most popular cognitive task used is the novel object test (Fig. 1a & b middle). This task is also very simple to implement and relies on an animal's innate curiosity and natural instinct to explore; it does not rely on the need to find a reward or escape from an aversive stimulus. This task typically consists of three separate phases; an animal is initially placed into an arena for a period of time, allowing it to explore and get used to the environment (habituation trial). Following this, the animal is allowed to explore the same arena which now contains a number of objects (training trial) - typically 2 objects. Then after a period of time, one of the objects is replaced with a novel one (test trial). The time spent by the animal on this trial exploring the novel object in comparison to the old one(s) is used as a measure of memory and recognition [56]. Deficits in this task have also been shown in a variety of AD models (for example [57] in scopolamine- and beta-amyloid peptide (Aβ)-induced amnesia models). The Y/T-maze is another popular test and consists of three arms and a single intersection. The task relies on the animal's natural tendency to alternate arms when exploring. It is commonly used to test spatial working memory as lack of alternation indicates a deficit, failing to recognise the previously visited arm [58].

Although contextual conditioning and passive avoidance tasks rely on the animal's ability to recognise and avoid a certain location due to the presence of an aversive stimulus, these tasks are used to examine long-term memory rather than as a test of anxiety (see discussion below). In the simplest version of the task, animals are allowed to freely explore two chambers on the first day of testing. On the second day, a shock is administered to the feet of the animal in one of the chambers. After a period of time (usually 24 h), the animal is placed into the chamber where no shock was given. If remembered, the animal will avoid entering the 'shocking' chamber and the time taken to enter this chamber is a good indicator of memory. Mostafa and colleagues [59] have shown that metformin ameliorated memory deficits on the passive avoidance task in their scopolamine-induced amnesia model of AD. Other tasks that examine cognitive flexibility, attention, executive functioning and episodic memory are rarely examined (see discussion
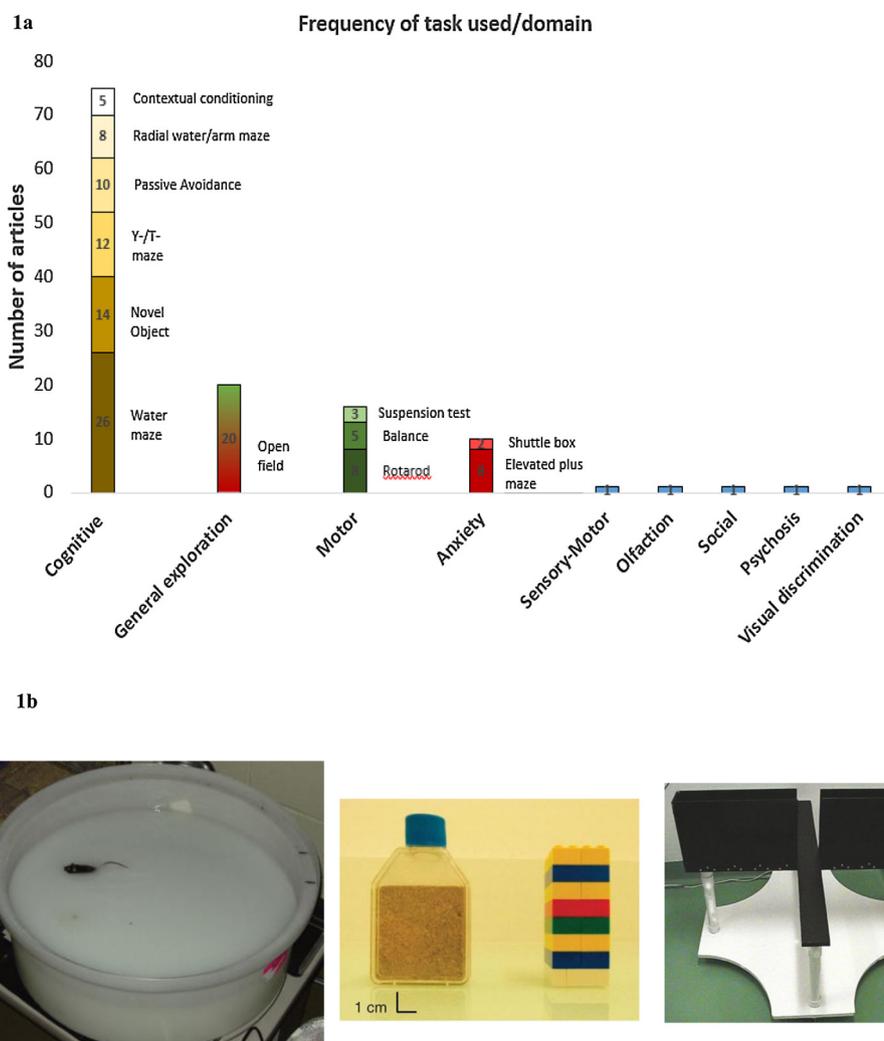
Fig. 1. a Frequency of tasks used in the articles sampled, separated by the domain being tested. **b:** Left. Photographs of the water maze (left) with hidden platform located below the water surface. Picture obtained with permission from Hamilton & Rhodes (2015) Prog Mol Biol Transl Sci. 135:381-406. Middle. A photograph of two typical objects (tower of Lego bricks and a flask filled with sand) used in an object exploration task. Picture obtained with permission from Leger et al. (2013) Nat Protoc. 8(12):2531-7. Right. Photograph of elevated plus maze. Picture obtained with permission from Walf & Frye (2007) Nat Protoc, 2, 322–328.

below), despite being critical clinical features of AD.

Within the *general exploration* domain, the open field task is very popular. Researchers in the majority of the articles examined (20 papers) used it to examine motor movement, anxiety, as well as, general activity. The task is very simple requiring little equipment and technology, which probably accounts for its popularity. In this task, animals are placed into an arena and monitored over a set period of time. The distance moved, time spent moving, the number of times the animal grooms or rears, and the time spent in various locations of the arena can all be measured. An animal that remains at the side of the arena and moves very little may indicate issues of motor and/or anxiety. The opposite may also be observed, such as, the hyperactivity found by Maroof et al. in their AβPPswe/PS1ΔE9 mice [60]. Other tests of general exploration and species-specific activity such as marble burying are seldom examined (see [61]), despite that such tasks may equate with various activities of everyday living measured in the clinic [62].

Motor functioning can be assessed more formally through a variety of tasks with the most popular of these, from our sample papers, being the rotarod, the balance beam and various tests of grip and suspension. The rotarod consists of a rod rotating on its longitudinal axis. The animal is placed onto the rod, which then slowly rotates. The rotation continues for a set period of time or, as in the accelerating version, the speed of rotation increases over the course of a few minutes. The time spent by

the animal on the rod before falling off is taken as a measure of balance. A simpler test of balance involves a narrow beam up to 100 cm in length (for mice, for rats a longer beam can be used) that is elevated and requires the animal to traverse its length within a set period of time. Some AD models, including the 5xFAD mouse model, show impaired balance and motor co-ordination, and these animals are quicker to fall off the balance beam in comparison to wild types [63]. Measuring the time that an animal can hang onto a grid or piece of wire while suspended at a height can be used as a good indicator of grip strength. Again, O'Leary [63] found that 5xFAD mice were faster to fall when compared to wild type (see also [64]). Other useful measures of motor impairment used with AD models include gait analysis (looking at shuffling behaviour & stride length, in 5xFAD mice [63]) and general monitoring of locomotion in an activity chamber (also with 5xFAD mice [65]).

Anxiety can also be tested by using a variety of tasks but the most popular one used from our sample (excluding the open field task) is the elevated plus maze (8 papers, Fig. 1a). The apparatus consists of two open arms (typically for mice, 30 cm in length and 5 cm wide) and two closed arms (with same length and width and 15 cm high) extending from a common central platform. The whole apparatus is elevated above the ground (see Fig. 1b right). The task relies on the animal's propensity to avoid open spaces. The animal is usually allowed a short period of time to explore the apparatus; the number of entries and the

time spent by the animal in the closed arms compared to the open arms can be examined. Although, Ardestani and colleagues [65] did not observe any deficits in 9-month old 5XFAD mice in the elevated plus maze, aged 3xTg-AD mice have been demonstrated to show a higher level of fear and anxiety using the same task [66]. Some researchers also use the light/dark box to examine issues relating to anxiety. The apparatus consists of two Perspex chambers; one chamber is brightly lit while the dark compartment is covered and dim. Both chambers are connected by a small opening at floor level. The amount of time the animal spends exploring the two chambers is noted, as are the number of light-dark transitions (a high number of transitions being indicative of lower anxiety). Again, anxious animals tend to avoid bright open spaces so will spend more time in the darker chamber; however, Plucińska & Crouch [67] found that their model animals remained significantly longer in the light chamber, suggesting reduced anxiety-related behaviour.

Other less common behavioural tasks that have been used with AD models include a *social discrimination* task (5XFAD mice [65]). During the first part of this task an unfamiliar mouse is placed in stainless steel wire grid cup in one part of the cage (an empty cup is placed on the opposite side). A test animal is free to explore around the entire area and can sniff and interact with the mouse in the cup. During the second phase, the mouse in the cup is moved to the opposite cup, while a new unfamiliar mouse is placed under the original cup. The time spent by the test mouse interacting with this new unfamiliar mouse compared to the previous mouse is a measure of social interaction and discrimination. One of the earliest signs of AD is impairment in olfactory functioning [68], yet very few animal models have examined this. One study by Xu et al. [69] showed that Tg2576 mice were impaired in odour memory but not with odour discrimination, suggesting that this might be a useful measure that could be extended to other models.

5. Adequacy of behavioural tasks

Behaviour by its very nature is multifaceted and complex, as such it is difficult to break down into constituent elements for any given situation. If we take memory as an example, there are different types (e.g. declarative, procedural, working, etc), which in turn can be sub-divided; declarative memories can be episodic or semantic in nature. Some memories can have both an episodic and semantic component; for example, spatial memory can have an episodic element (e.g. recalling a location wherein an event took place) and a semantic element (e.g. my house is located beside the park). Memories can rely on multiple senses (olfactory, visual, auditory, etc.); they can span time – from a few minutes to a life time; they can have different storage capacities, and they can rely on different neural structures. A single task, therefore, can not capture all elements of a particular behaviour. Even if we were to try and capture most aspects of the behaviour (such as memory) we would be overwhelmed by the amount of tasks we would need to employ - even if such tasks existed. Despite this seemingly pessimistic view, many ingenious tasks have been used very successfully to capture a wide-variety of behaviours and sub-behaviours, and these have also been applied to various models of neurodegeneration including AD. In the following sections we discuss the various tasks used in AD animal models but highlight how some of them are open to interpretational, translational and procedural issues.

5.1. Interpretational issues

One of the difficulties with some commonly-used behavioural tests is that they are open to interpretation, it is often difficult to know what exact behaviour is being examined by a particular task. Furthermore, different authors use the same task to measure different behaviours. For example, the Open Field test (described above) is a simple, easily administered test, in which the animal is free to explore an open environment for a specified period of time. The animal can easily be

tracked using video-tracking software or certain behaviours can be manually recorded by the experimenter. The test has been used for a range of behaviours including the examination of general exploration and activity, as well as, specifically allowing an animal to habituate to an environment, before the introduction of objects (as in the case of the object recognition task [70]). More often, especially in examining models of AD, the task has been used to examine both anxiety [71] and locomotion [63] behaviours, as well as general exploration. The time spent by the animal moving/not moving, the time spent by the animal in certain parts of the environment (e.g. at the side or in the open), as well as changes in other natural behaviours (times rearing, grooming) can be indicative of a deficit in the model but the difficulty is defining the precise nature of this deficit.

The Passive Avoidance task is another important test in AD models and is traditionally used to examine learning and long-term memory [58]. In the task animals are required to avoid an environment in which an aversive stimulus (e.g. a shock) has been previously administered. Typically, the task is used in AD models to assess deficits in mnemonic functioning [72]. However, passive-avoidance is also a fear-motivated task [73]. Animals are conflicted to avoid an environment that they would naturally prefer (e.g. a dark chamber). This approach-avoidance choice relies not only on memory but motivation and fear too. Furthermore, to complicate matters many authors use the passive-avoidance task as a test of non-spatial memory [74]; this is despite the important role of context and location in the task. Other researchers use the task to test memories at various intervals (e.g. at 24 h [75] or 48 h and 1 week retention [66]). Given that different memories recruit different structures across time [76] and that some memory systems are quickly impaired, while others remain spared, careful consideration should be given to why and how a particular task, like the passive-avoidance test is used. Furthermore, given the analogy between passive avoidance and the Claparède phenomenon in AD patients (whereby an aversive event could be recalled even if the patient's short-term memory was impaired), caution should be exercised if considering the passive avoidance test [77].

Spatial memory deficits are key characteristics of AD [34]. Given that brain areas such as the entorhinal cortex, hippocampus and anterodorsal thalamus are heavily involved in spatial navigation [78,79], as well as, being some of the first structures affected by AD pathophysiology [80], it is not surprising that many authors recently have suggested that spatial navigation deficits have the potential to detect underlying pathophysiology in Mild Cognitive Impaired (MCI) and pre-clinical AD [34]. The water maze task, which requires animals to search for a hidden platform located in a pool of water by using landmarks in the environment, is considered the 'gold standard' for spatial navigation and memory [45]. It is a task that depends on an intact hippocampus [49,50] and has been used as the main cognitive test in AD models in the majority of the papers that we analysed (for example [58,59]). Aside from the procedural issues of the water maze procedure (discussed below), most authors do not consider adequate control measures for behaviour in this well-validated task, and any results obtained therefore may be open to interpretation. For example, if an AD animal model does not find the hidden platform as quickly as the wild-type, does this automatically suggest that the model mouse has impaired spatial memory? The water maze task relies on many behaviours aside from spatial memory – animals must be able visualise external cues, they must be able to swim normally, they must be motivated to find the platform and not be too distressed. Age [81], motor and visual impairments [82] and stress [83] can all lead to a poor water maze performance; many of these aspects are not directly related to AD, are not the primary behaviour of interest by the researcher but yet may be a characteristic of the particular model used. Work by Spencer and colleagues, for example, showed that the cause of failure in the Morris water maze by aged rats, was due to retinal degeneration [84]. One possible way to control for this would be to use cue training to exclude affected rodents [85]. Hyperactivity and hypoactivity have also been

reported in a number of models including APP$_{695}$SWE (Tg2576) and APP$_{770}$SWE, respectively [86–88]. In some models, hyperactivity has been shown to directly affect water maze performance. For example, Baeta-Corral & Giménez-Llort [51] showed that escape latency, distance and speed in finding the hidden platform were all affected by hyperactivity in 3xTg-AD mice. A recent simple water maze control task developed by Barry & Commins [89] may prove a useful method to separate spatial learning from other characteristics, including motor activity and stress in AD models, and aid in the interpretation of any deficits observed.

5.2. Translational issues

Clearly the goal of any animal disease model is to try and match it, as closely as possible, to what occurs in humans. Götz and colleagues [30] offer a nice account of what to expect in a good animal model for AD. In sum, the authors suggest that models should express the main cellular features of AD, including neurodegeneration, plaques, neurofibrillary tangles, and changes along specific molecular pathways. These histopathological features should be initially confined to the specific brain regions before spreading to other brain regions, mapping onto the progression in humans. Furthermore, any cognitive symptoms observed in humans, which are many, should been also seen in the animal model, and the occurrence of such deficits should be again linked to the progression of the disease. The question asked here is whether the behavioural tasks used in animal models and the cognitive functions that they purportedly test are matched with those used in the clinic and the deficits observed in patients.

First, many tasks used in AD models may not be directly matched to tests carried out in the clinic. The passive avoidance task is a good test for long-term memory in animals; however, can this test be directly compared to long-term memory or other memory tests typically used in the clinic? Avoiding a particular location as a result of a single fearful occurrence may not be a common occurrence among those with AD. Given ethical and other considerations, this paradigm may not be appropriate to use as a test with a patient sample. The novel object recognition task has also been used in many animal models, primarily to examine both short and long-term object memory [70]. Although object recognition relies on structures of the medial temporal lobe (including the perirhinal cortex [90] and the hippocampus [91]) and is impaired in AD [92], some studies have reported that AD patients are as accurate as controls in object detection [93]. In the clinic, memory is often tested initially through a variety of general tests of cognitive impairment including the Mini-Mental State Examination (MMSE) and the Montreal Cognitive Assessment (MoCA), both of which seem to be effective for detecting AD [94]. More specific tests can be used during follow-up to determine the type and range of memory issues. For example, short-term memory can be assessed by using the Digit Span, whereby patients are required to recall an increasing sequence of numbers backwards or forwards, while long-term memory can be assessed by adapting the California Verbal Learning Test (CVLT [95]) or the Philadelphia Verbal Learning Test (PVLT [96]) by including a 20-minute delay [97]. More specific tests may include memory tests for faces [98] and paired associate learning [99]. Therefore, there is a mis-match between how some important functions (e.g. memory) are examined in the clinic compared to the laboratory. Often clinicians rely on tests that are easy-to-use, easy-to-administer and are quick to complete. This is in contrast with laboratory tasks that often require multiple trials administered across hours or days.

Second, the difficulty in translating cognitive domains tested in the clinic into behavioural animal tasks is that many screening tools, as expected, rely on distinctively human characteristics and are therefore difficult to translate directly to animal (particularly rodent) models. Such domains include language (comprehension, writing, naming and reading, e.g. tested as part of Addenbrooke's Cognitive Examination (ACE-III) [100], verbal fluency [101], abstract thinking (e.g. in the Brief

Kingston Standardised Cognitive Assessment Revised (BKSCA-R [102]) and copy and construction (e.g. in the Short Test of Mental Status (STMS) [103],). However, there are a number of domains tested by the various screening tests that can translate to animals but, unfortunately these are seldom used as behavioural tasks with AD models. For example, one of the most translatable tests for AD, and one that is considered the gold standard for memory in the clinic, is tests of episodic memory [104,34]. As mentioned, episodic memory refers to the ability to remember particular events in one's life. For example, you might remember attending a particular birthday party when you were young. Episodic memory deficits are a common feature of AD and is considered a sensitive cognitive measure of underlying pathophysiology, particularly as episodic memory is thought to rely on a functioning hippocampus [105]. Despite tests of episodic memory being used as a diagnostic tool in the clinic, very few studies in animal models attempt to model this, despite the existence of a number of analogous tasks. One such test is the What-Where-Which (WWWhich) developed by Eacott & Norman [106] and based on the work by Clayton & Dickinson [107]. In this task animals, spontaneously explore an environment and attempt to associate an object (What), its location (Where) and the environmental context (Which). In another version, the WWWhen task, animals have to associate an object (What), its location (Where) and the temporal context (first or second occurrence – When [108]). Clear deficits in episodic-like memory have been found in a number of different models of AD that have used such tasks (e.g. Tg2576 APPswe [109],) and the deficits observed seem to map to the progression of the underlying pathophysiology (3xTgAD [110],).

Attentional and executive function deficits are other key features in patients with AD and form part of the diagnosis criteria in the Diagnostic and Statistical Manual of Mental Disorders (DSM–5). There are a number of different tests that can be used for modelling attention and executive functioning in animals, but again these are seldom used. For example, one popular test is the five choice serial reaction time task (5-CSRTT). This task was originally developed by Carli and colleagues [111] but has been modified and adapted over the subsequent years. In the task, the animal must detect and respond to brief light stimulus emanating from an aperture (from a potential range of 5–9 openings) in an operant box. Sustained attention can be measured by examining when the animal responds to the incorrect hole versus the correct hole from where the light appeared (errors of commission) or fails to respond within a certain time (errors of omission). Other measures of response inhibition including premature responding or continuing to repeat previous mistakes (perseveration) may also be used to test executive functioning [39]. The task is similar to the Continuous Performance test in humans [112], whereby participant are required to respond only to a pre-specified target among a continuous sequence of stimuli, or the alternative Sustained Attention to Response Task (SART [113]), where participants are required to respond to every stimuli but withhold their response on a pre-specified target. Despite being a robust and valid task of attention, the 5-CSRTT is seldom used with animal models of AD; however, the few that have used it, including Romberg [114] with 3XTg-AD mice, have found performance deficits. However, Bharmal and colleagues [115] did not find any difference in performance at 5 months using TgTau-P301L mice, when compared to wild-type. Although, the WWW, 5-CSRTT and other tasks are innovative and are possibly provide measures that are analogous to some of cognitive deficits observed in human patients, caution must always be applied before direct comparisons can be made.

Third, the reverse could also be considered; some tasks used in animal models may be applicable in the clinic, but seldom form part of the screening or assessment process. For example, although AD is considered mainly as a cognitive disorder, AD patients can develop motor deficits, which may put patients at an increased risk of falling [116]. Many rodent models do include tests of motor functioning, such as, tests of balance, grip and strength. Motor impairments in the animal models, including weakening grip strength (3xTg-AD [117],), balance

issues and slower speed (5xFAD [63],), often match the motor deficits observed in patients [118]. Recent evidence further highlights the link between gait, balance and fall issues among those with AD [119] and older adults, with a particular emphasis on the dual cognitive-motor aspect of walking and motor co-ordination [120]. For example, there is an increased risk of falling or gait changes if patients are required to perform a cognitive task while walking (dual-task paradigm – [121]). Simple gait analysis, dual-task measures and other tests of motor functioning may be easily adopted in the clinic and offer a promising approach for the detection and monitoring of AD progression.

Fourth, recent research has highlighted other potential behaviours that are typically not considered in either animal models of AD or the clinic. Some of these behaviours could be examined with relative ease. For example, AD patients have been shown to have disturbed sleep patterns and circadian rhythm deficits compared to healthy older adults. Specifically, AD patients tend to show a reduced amplitude of circadian rhythmicity, show a more fragmented sleep pattern, wake with greater frequency during the night and sleep more during the day [122]. Although, there are conflicting results on the exact nature of the deficits, which seem to depend on the type and severity of cognitive impairment [123], most studies find greater circadian deficits in cases of severe AD and less impairments with MCI and milder forms of AD [124]. Although there have been very few studies examining clock gene expression in AD patients, Cermakian and co-workers [125] found difference in *PER1*, *PER2* and *BMAL1* gene expression in post-mortem brain tissue of AD patients compared to controls. Similarly, examining sleep and circadian patterns and the underlying molecular expression are rarely done in AD animal models; such behavioural measures could easily be incorporated when testing. For example, using a simple running wheel, Stover [117] monitored the number of rotations that 3xTg-AD mice completed in the wheel over a 7 day period and reported a disrupted circadian rhythm compared to wild-type controls. While Plucińska and colleagues [67] used a home-cage monitoring system to examine distance moved in hBACE1 mice. Furthermore, it would be very useful to correlate circadian and sleep disturbances with cognitive measures, as well as, the underlying pathology.

Other behaviours, such as, depressive symptoms have long been observed in patients with AD [126]. Despite this, there has been very little examination of depression-like symptoms in animal AD models. Tests such as the Tail Suspension Test or the Forced Swim Test [127] may be easily incorporated as extra behavioural measures. In the Tail Suspension Test, animals are usually suspended from a horizontal bar by their tails from a certain height and for a specific time length (e.g. ˜120 cm above a table for 6 min [128]), time spent immobile during this period is used to measure depressive-like symptoms. Likewise, for forced swim test animals are placed into a small container of water (from which they cannot escape) for a certain period (e.g. 6 min. [127],). The amount time that the animal stops swimming (i.e. immobile) is taken as the animal having 'given-up' and an indicator of depression-like behaviour. Very few studies have used such tasks in AD animal models, but using the tail suspension test, Iascone and colleagues [129] showed that hAPP mice exhibited depressive behaviour at 13–15 months but not at 5–7 months.

5.3. Procedural issues

Many of the same tests are used very differently, despite the fact that they are supposedly testing the same function. Using a non-standardised testing apparatus and procedures make it very difficult for researchers to compare results, to allow for replication of experiments and the development of norms. This is particularly important if we are trying to compare various models of AD. Using different models with different procedures and/or a different testing apparatus, how can findings be generalised across multiple AD models and then translated to humans and the clinic? In Table 1 we highlight how different labs use the same task. This is not meant to be an exhaustive list but merely used

to point out discrepancies between both the test apparatus (e.g. the size, shape, physical dimensions) and the procedure used (e.g. number of trials, inter-trial-interval length, trial length etc).

From our sample of the literature, the most consistent behavioural task used, both in terms of the apparatus and procedure, is the elevated plus maze. The time allowed for mice to explore the maze is generally 5 min (with just one of our studies giving 8 min, [57]). Likewise, the apparatus used is very standard with a typical arm length, width and height (of closed arms) of 30 cm, 5 cm and 15 cm, respectively. The only major difference between the various studies is in terms of the height that the apparatus is elevated, ranging from 50 cm to 100 cm. This may be an issue as 'elevation' and an animal's fear of open spaces (including the third dimension) are critical to the task itself.

The next task that remains relatively consistent between research groups is the water maze task. All experiments use a circular pool; while the diameter of the pool does range between 100 cm–180 cm, most of our sample experiments have a pool diameter of 120 cm. Similarly, although the size of the platform ranges between 7 cm–17 cm in diameter, the relative platform size to the overall pool size is approximately 8% (although see [130] at 12%). Note, that these values are used typically with mice. For rats, the pool is often 180–200 cm s in diameter (see [47] for example). Size is critically important as there have been reports to demonstrate that if the pool is too small relative to the goal size, rats are able to solve the water maze without the use of spatial cues [131]. In terms of procedures, the maximum time allowed for a given trial is typically 60 s (although both O'Leary [63] and Plucinska [67] allowed 90 s for their trials). Where the biggest difference between the experiments arise is in terms of the training procedure; animals can receive anywhere between 10 trials [132] to 30 trials (6 trials/day for 5 days [130],). Most authors tend to give 4 trials/day for 4 or 5 days (i.e. 16–20 trials, see Table 1). The worrying aspects about some of the water maze procedures used is that longer trials or multiple trials with a short inter-trial interval, can lead to reduced swim speeds, most often in mice and aged rats. While it is possible this is due to fatigue, the more likely explanation is hypothermia due to the short inter-trial interval [133,134]

From an examination of our sample of studies, the rotarod test for motor co-ordination is very consistent in terms of the apparatus used, possibly because rotarods are commercially available and are fairly standardised. However, the manner in which it is used differs across research groups. For example using mice, the baseline rotation rate can range from 0 rotations per minute (RPM, i.e. not rotating) to 4 RPM, while the maximum rotations can range from 40 to 48 RPM. The time allowed to reach the maximum rotation speed also differs from 5 to 10 min (see Table 1, first column). Similarly, many authors simply use a single trial [54], while others train animals across trials and days [67]. In a similar fashion, the balance beam also tends to be consistent in terms of the apparatus used – most authors using mice have a beam of approximately 80–100 cm in length, 2–3 cm in width and 30–50 cm s in terms of elevation. However, the training procedure differs across experiments, with some authors using a single 2-minute trial [117], while others use multiple training trials across days [130]. At the more inconsistent end for both apparatus and procedure used are the T/Y maze task and the novel object recognition task. For example, Davis [135] used a T-maze with arms of 19 cm in length and 4.5 cm wide and 9.5 cm high walls (with mice). Whereas, Olmos-Alonso [136] used arms of $30 \times 10 \times 29$ cm. Similarly, some authors simply allow animals to explore the maze for a set period [67]; while others train across multiple trials and days [136]. The novel object task also varies in its procedures - from allowing animals to habituate across days [70] to a single trial [65], from separating training and testing sessions by minutes [61], an hour [54] or 24 h [65]. In general, two items are used during testing, one of which is then replaced. Authors seldom describe the items; responses to plastic or more natural objects (e.g. stone or wood) may produce very different reactions and these are not usually considered. Finally, despite being one of the most popular tests used for

Table 1
Comparison of the different apparatus and procedures used by different laboratories when performing the 'same' test on laboratory animals.

Tests of Motor Functioning			Tests of Cognitive Functioning			Tests of Anxiety		
Rotarod (balance)			Water maze task (spatial memory) -mice			Elevated Plus maze		
Study	**Environment/ Apparatus**	**Procedure**	**Study**	**Environment/ Apparatus**	**Procedure**	**Study**	**Environment/ Apparatus**	**Procedure**
Capurro et al. 2013		2.5 RPM baseline and reaching 42.5 RPM over 5 min (increasing speed every 30 seconds).	Capurro et al. 2013	Circular pool diameter, 120 cm; height, 62 cm). Platform (diameter 10 cm, height 31 cm).	2 daily sessions comprising 4 consecutive 60-second trials, 15-second inter-trial interval.	Ardestani et al., 2017	The apparatus consisted of two open arms (30 cm long × 5cm wide) and two closed arms (30 × 5 × 15 cm) and to a height of 50 cm above the floor.	Allowed 8 mins to explore.
Young et al 2016	Rotating, white, ridged rubber rod, separated by fixed, white polyacrylate partitions into five, 6 cm-wide sections.	3 trials in total, consisting of two runs each, over three days. Rod rotated at a baseline speed of 4 RPM. reaching 40 RPM over 10 mins.	Li et al. 2017	Circular pool 120 cm in diameter	4 sessions, with one session (four trials X 60s) per day. 30 s on platform.	Sabogal-Guáqueta et al. 2015	The apparatus consisted of two open arms (30 × 5 × 0.25 cm) and two closed arms (30 × 5 × 15 cm) extending from a common central platform (5 × 5 cm), and to a height of 60 cm above the floor.	Allowed 5 mins to explore.
Plucińska et al. 2014		Testing consisted of four trials/day for 2 consecutive days with ITI of 2–3 min. Rod accelerated from 1 to 45 RPM over 5 mins.	Ardestani et al., 2017	Large water tank (178 cm in diameter). Platform of 17 cm diameter circular	4–6 days of hidden platform training (60 s). Each day consisted of 4 trials (30 min intertrial intervals). Rest on platform for 10s.	Young et al., 2016	The apparatus consisted of two open arms (35 × 5 cm) and two closed arms (30 × 5 × 12 cm) and to a height of 100 cm above the floor.	Allowed 5 mins to explore.
Webster SJ, Bachstetter AD 2013	Rotating rod (3.18 cm diameter) in lanes 11.5 cm wide.	Rod accelerated from 0 to 40 RPM over 5 mins.	Wang et al 2017	Circular pool (diameter: 120 cm, height: 50 cm).	The mice were trained four times per day (60 s/trial) for 5 days. Rest on platform for 10s.	Meng et al. 2017	The apparatus consisted of two open arms (30 × 6 cm) and two closed arms (30 × 6 × 15 cm) extending from a common central platform (6 × 6 cm), and to a height of 80 cm above the floor.	Allowed 5 mins to explore.
O'Leary et al., 2018	Rotating rod (3 cm diameter), separated into four 11 cm	Mice completed 3 days of training with 6	Hochgräfe et al., 2015	180 cm circular pool. Platform 15 cm round	Each mouse performed 4 swimming trials per day	Plucińska et al. 2014	The apparatus consisted of two open arms (35 × 5 cm) and two closed arms (35	Allowed 5 mins to explore.

Table 1 (*continued*)

	sections by circular Plexiglas dividers (15 cm high).	trials per day. Rod accelerated from 0 to 48 RPM over 6 mins.			(maximum duration 90s, 10 min inter-trial interval) for five consecutive days.		× 5 cm) extending from a common central platform (5 × 5 cm), and to a height of 42 cm above the floor.	
Stover KR, Campbell MA 2015	Rotating rod (3 cm diameter) with four opaque Plexiglas barriers (15 cm diameter) dividing the rod into four 11 cm sections.	Rod accelerated from 0 to 48 RPM over 6 mins.	D'Souza et al., 2015	Circular pool 1.2m in diameter. (Platform15cm in diameter)	5 day hidden trial block (6 trials per day). 60s allowed for each trial.	Lok K, Zhao et al 2013	The apparatus consisted of two open arms (30 cm long × 5cm wide) and two closed arms (30 × 5 × 15 cm) and to a height of 50 cm above the floor.	Allowed 5 mins to explore.
			Sabogal-Guáqueta et al. 2015	Pool 100cm in diameter and 30 cm in height. Platform (7 cm diameter)	Ten sessions or trials were performed, two complete sessions per day for five days Each session consisted of four successive subtrials.			
			Plucińska et al. 2014	Pool 150cm in diameter and 50 cm in height.	Four trials each day (ITI: 30 min) – 4 days. 90s per trial max.			
Balance beam (balance)			**T/Y maze**			**Open Field**		
D'Souza et al., 2015	A beam 1.8 cm wide, 100cm long elevated by 50cm. An empty cage with nesting material served as a finish point at the end of the beam to attract the mouse.	3 Days in total. Day 1 and day 2 were training days in which the mouse was encouraged to cross the beam 3 times with a 10 minute break between each trial. On day 3, the mice were given 3 trials, and the two	Ardestani et al., 2017	3 arms separated by 120 ° angles. Each arm 40 cm long X 8 cm wide X 15 cm high walls.	Allowed 5 mins to explore.	Capurro et al. 2013	Plastic square box 46.5 × 43.5 cm	Allowed 15 mins to explore.

Table 1 (*continued*)

		best times required to cross the beam and enter the finish box were averaged.						
Webster SJ, Bachstetter AD 2013	A beam 3 cm wide, 80 cm long, elevated by 30 cm to reach an enclosed goal box.	Mice were placed on the beam at one end and allowed to traverse the beam to reach the goal box. This was repeated using decreasing size (3 cm, 2 cm, and 1 cm).	D'Souza et al., 2015	Not stated	Allowed 10 mins to explore.	Sterniczuk et al., 2010	White, plastic, circular pool, 123 cm in diameter and 36 cm deep	Allowed 5 mins to explore.
O'Leary TP, Robertson A, 2018	A beam 2 cm wide, 100 cm length, elevated by 40 cm.	Mice allowed to move along the beam for 120 s.	Young et al., 2016	3 arms separated by 120 ° angles. Each arm 44 cm long X 7 cm wide X 25 cm high walls.	Allowed 5 mins to explore.	Ardestani et al., 2017	Square arena (76 cm × 76 cm x 50 cm)	Allowed 10 mins to explore
Stover KR, Campbell MA 2015	A beam 2 cm wide, 100 cm length, elevated by 40 cm.	A single trial lasting a maximum of two minutes was given	Olmos-Alonso et al 2016	T- maze with 30 × 10 × 29 cm arms,	30 s to facilitate habituation. four times a day for 5 days (a total of 20 trials) with an average spacing of 2 h in between each trial.	Hochgräfe et al 2015	A 50 x 50 cm arena	Allowed 15 mins to explore
			Davis KE, Burnett K, Gigg et al. 2017	T-maze with arm 19 cm long X 4.5 cm wide X 9.5 cm high walls.	Forced choice spontaneous alternation behaviour. 5 min acquisition phase, an inter-trial-interval (ITI; 2 or 30 min) and a final 5 min test phase. ITIs of 2 min and 30 min	Young et al 2016	A 35 × 30 × 30 cm arena	Allowed 30 mins to explore.
			Plucińska et al. 2014	Each arm 60 cm long X 10 cm wide X 10 cm high walls.	Allowed 10 mins to explore	Olmos-Alonso et al 2016	A cage 27 × 27 × 0.3 cm	Allowed 5 mins to explore.
						Meng FT et al. 2017	A wooden box (50 × 50 cm x 25	Allowed 5 mins to

Table 1 (continued)

							cm height)	explore.
				Novel Object			**Light/Dark box**	
			Li et al. 2017	Open field box (50×50×25 cm)	Habituated for 15 min every day for 3 days. Two novel objects placed in the centre. In the 1h and 24h retention tests, mice were placed back into the same box for 15 min, and one of the objects was replaced with a novel object.	Meng FT, Zhao J, 2017	One light box and one dark box, and both boxes were the same size (25 cm × 25 cm × 25 cm),	Allowed 5 mins to explore.
			Ardestani et al., 2017	20 × 40 cm arena	Habituated for 15 min the day prior to testing. Testing: given 2 identical unfamiliar objects positioned 5 cm away from the walls. Allowed 10 mins to explore. Twenty-four hours later, one of the objects was changed for a new unfamiliar one. Allowed to explore for 10 mins.	Plucińska K, Crouch B 2014	One light box (30 × 30 × 30 cm) brightly illuminated and a smaller dark compartment (30 × 20 × 25 cm) covered with Plexiglas lid.	Allowed 10 mins to explore.
			D'Souza et al., 2015	Not given	Habituated for for 3 days. Two novel objects placed in the centre.	Lok K, Zhao H, 2013	**Shuttle Box** 50 × 25 × 25 cm two way automated In each trial, a tone (60 dB) was used as the conditioned stimulus (CS), and was presented for 5 s. Subsequently, an electrical foot shock (0.4 mA), used as the unconditioned stimulus (UCS) was given for 10 s. If the mouse moved	

Table 1 (*continued*)

			Reference	Apparatus	Procedure		Procedure	
					On the testing day, animals were allowed to explore the objects until they accumulated 30 seconds of total object exploration time.		to the other side of the box during the period of CS alone (i.e., a latency of less than 5 s), no shock was presented and the response was scored as an active avoidance. If the mouse did not cross to the other side of the box during the first 5 s of the CS, a foot shock was delivered until the mouse escaped or until 10 s had elapsed. Mice completed 4 days of acquisition training with 50 trials per day.	
			Young et al 2016	35 cm × 30 cm × 30 cm height	During habituation, mice were presented with two identical objects for six minutes. Ten minutes later, one object was moved for 6 mins. Twenty minutes later, one object was replaced with novel one for 6 mins.			
			Webster SJ, Bachstetter AD 2013	Open field box 39.4 cm × 78.8 cm	Habituated for 15 min the day prior to testing. Allowed 10 mins to explore 2 identical objects. One hour later, one of the objects was changed for a new unfamiliar one. Allowed to explore for 10 mins.			

AD models (see Fig. 1a), the open field test is the most inconsistent in our sample. For example using mice, the apparatus can range from a square box [65], a circular arena [66] to a home cage [136], as such, the size of the arena changes dramatically (e.g. 50 × 50 cm, 35 × 30 cm, 76 × 76 cm, 123 cm in diameter and 46 × 43 cm, see Table 1). In addition, the time allowed to explore the arena varies from 5, 10, 15 and up to 30 min; this is often irrespective of the size of the arena. Such inconsistency makes direct comparisons between AD models very difficult, especially if such an AD model displays hyperactivity (see discussion above).

6. Recommendations

Although animal models play an important role in understanding the mechanisms of a neurological disorder, it is clear from the discussion above that the choice and adequacy of behavioural tests used is critical to allow the interpretation of results. We detail below some recommendations and potential developments in the field of behavioural testing in models of Alzheimer's disease, that are also applicable to many other neurodegenerative disorders.

6.1. Better behavioural controls and experimental design

While there are guidelines and procedures for conducting pre-clinical testing, they are not always complied with. This has led to considerable criticism of animal models as the largely positive results from pre-clinical testing have not translated into clinical results and this needs to be improved. Strict use of the existing animal model guidelines, such as ARRIVE ('Animal Research: Reporting of In Vivo Experiments'), is strongly recommended and vital to improve the translatability of in vivo experiments [137]. Good experimental design, including randomisation and blind testing is critical; however, it is not always done, or reported in publications [138]. For example, the use of blinding of experimenters to treatment group allocation is important to ensure that in all tests, including behavioural assessment, there is no bias which could affect the outcome of the tests, including the well-characterised experimenter effects illustrated by Rosenthal [139]. However, as with randomisation, this is not always clearly described in the methods. This was clearly highlighted by Egan and colleagues who carried out a meta-analysis of interventions in transgenic mouse models of Alzheimer's disease. Amongst others within this study, they showed

Table 2
Suggested dimensions and procedures for various tasks.

Task	Species	Apparatus dimensions (approx.)	Procedure	Sample references
Motor				
Rotarod	Mice	Rod of 3 cm in diameter, supported 30 cm above the base. The surface of rod is knurled in parallel ridges to help grip. 2 flanges (20–30 cm in diameter) or more (to accommodate multiple animals to be tested simultaneously) are separated by a distance of 3–6 cm.	**Habituation period:** 1 min trial on static rotor. After 15 min break, another 1 min trial at 4 r.p.m. **To test motor performance:** A single 300 s trial with rotor accelerating from 4 r.p.m to 40 r.p.m is used. Latency to fall is recorded. **For motor learning:** 3 × 300 s trials with ITI of 15 min. Rotor increasing from 4 to 40 r.p.m during each trial.	[141] [142]
	Rats	Rod of 7–8 cm in diameter.	As above	[143]
Balance Beam	Mice	Wooden beam of 80-100 cm in length, 1-2 in width and 30–50 cms in terms of elevation.	**Training** Mice given 3 trials to cross beam. With 10 min break between each. Time taken to cross beam is measured. **Test** Average time to cross beam taken from 2 successful trials.	[144]
	Rats	Wooden beam of 100-120 cm in length, 2-3 cm in width and 50 cm elevation.	As above	[145]
Cognition				
Water maze (Spatial acquisition)	Mice	˜120 cm diameter circular arena with height of 50 cm. Circular or square platform of 10–12 cm in diameter or 10 cm^2, respectively. Located 0.5–1 cm below water surface. Search area:target ratio of ˜117:1.	Allow 1 min for each trial. Inter-trial interval of 15 s. 4 trials per day for 5 or 6 days. (20–24 trials). Latency to find hidden platform is measured.	[146]
	Rats	˜200 cm diameter circular arena with height of 50 cm. Circular or square platform of 10–12 cm in diameter or 10 cm^2, respectively. Located 1-2 cm below water surface. Search area: target ratio of ˜346:1.	Allow 1 or 2 min for each trial. Inter-trial interval of 15 s. 4 trials per day for 5 or 6 days. (20–24 trials). Latency to find hidden platform is measured.	
T/Y maze (Spontaneous alternation)	Mice	Start alley: 30 cm length × 10 cm width. Goal arm (X2) 30 cm length × 10 cm width. Wall height 20 cm	**Continuous trial procedure** Allow min 5 min. to explore Sequence of arms is recorded **Discrete trial procedure** Place animal in start arm and allow to choose a goal arm. Confine animal in goal arm for 30 s. Replace animal in start arm and allow animal to pick from the two goal arms. Number of correct choices noted. Each trial should last < 2 min. 2/3 trials across multiple days is suggested.	[147]
	Rats	Start alley: 50 cm length × 16 cm width. Goal arm (X2) 50 cm length × 10 cm width. Wall height 30+ cm	As above	
Novel Object	Mice/rats	There is no standard size or shape but a square or rectangle apparatus of dimensions between 30 cm × 30 cm × 30 cm and 50 cm × 50 cm × 50 cm are typically used. 2 objects (key feature is to maximise the difference between objects without inducing a strong preference for one). Objects of different shape, texture, size and brightness can be used. Rodents have a difficulty in discriminating colour, so this should be avoided.	**Habituation** Allow animal to explore empty arena for 5 min. **Familiarisation** 24 hours later, allow animal 5 min to explore arena that now contains the two objects, located at 5 cm from wall. Time exploring each and both objects is recorded. **Test** Replace the two familiar objects, one with a triplicate copy (to ensure no olfactory residue) and the other by novel object. Allow 5 min to explore.	[148]
Anxiety				
Elevated plus maze	Mice	4 arms (two open without walls and two enclosed by 15 cm high walls). Each arm 30 cm long and 5 cm wide. Elevation 40 cm above ground	5 min exploration. Time and Number of entries into open and closed arms should be recorded.	[149]
	Rats	4 arms (two open without walls and two enclosed by 30 cm high walls). Each arm 50 cm long and 10 cm wide. Elevation 50 cm above ground.	5 min exploration. Time and Number of entries into open and closed arms should be recorded.	

that of the publications they examined, less than 25% reported blinding of experimenters [7]. In the same study, it was demonstrated that when randomisation and blinding were conducted in studies, these led to smaller effect sizes. Careful sample size calculations should also be made to ensure the experiment is sufficiently powered to detect any difference (also an element that is often overlooked [7]).

Linked to good experimental design is having adequate behavioural controls. While most studies use wild-type or a saline condition as a control for various animal model, researchers seldom if ever control for behaviour or highlight the precise behaviour they wish to examine. Many behavioural tasks tap into a range of functions, deficits in any one of these functions may be misinterpreted as a feature of the diseased model. For example, an animal that does poorly on the water maze task does not necessarily imply spatial learning deficits. Have motor, visual, motivation, stress and other factors been completed ruled out? If an AD model tends to show motor deficits, does a lack of movement in the

exploration task or slowness in the water maze task reflect a lack of motivation or learning, respectively? Researchers need to be aware of the limitations of the various tasks and use appropriate behavioural controls (see [89]).

6.2. Standardisation of testing

One of the major aspects that we highlighted in this review was the lack of procedural consistency in many of the behavioural tasks currently in use. If different research groups are using different procedures (even if the apparatus is similar), it makes comparison and replication of results very difficult. There are some excellent reviews and recommendations of protocols currently available in the literature and these match closely with the consistently used procedures identified in our sample. In Table 2 we provide a summary of some of these protocols and appropriate references for more detailed descriptions. We can not describe all tasks or the range of different protocols that could be used to examine various functions, even using the same tool (e.g. spatial working memory, reference memory, procedural memory and cued memory can be all tested in the water maze task by using different procedures). However, at a minimum we recommend researchers to adhere to such published and well validated protocols that have been consistently shown to work across various AD models.

6.3. Interaction between stakeholders

Public-patient involvement (PPI) is increasingly recognised as an important aspect of research, especially human work. Here patients/public work closely with the researcher to help set priorities and design experiments. Research is said to be done 'with' patients or members of the public rather than 'for' them. While a full PPI may not be appropriate with some animal work, researchers using animal models should link in with patient groups and interact with clinical practitioners as much as possible. This will help researchers to understand behavioural, cognitive and other issues in practice and may help to overcome some of the translational issues discussed above. Furthermore, while some of the animal models used are relevant to the clinic, many are not. Therefore, there is a need to re-evaluate the models and behavioural tests used to ensure they are appropriate and relevant. For example, as described above, there are many behavioural tests that are relevant to the clinical condition, but they are often the least popular measures used. Consequently, more careful consideration and planning of phenotypic measurement is required. It is important to note that interactions between patients, clinicians and researchers should not be unidirectional but learning should be reciprocated. For example, although spatial navigation impairments are observed in patients and in animal models [34], these deficits are currently not routinely tested in the clinic. This can be achieved with greater interaction between all stakeholders.

6.4. Innovative in design of tests

Acknowledging that there is often a difficulty in directly translating many uniquely human tests to animals, particularly those with a strong language component, researchers need to be innovative in their approach. Animal tasks such as the WWWhen task or the five choice serial reaction time task are excellent examples of directly translating episodic memory and attention deficits observed in AD patients. However, caution is still required as any novel test would need to be appropriately validated. One approach to doing this with models of any of the neurodegenerative conditions would be to test the model with a therapeutic agent that has been demonstrated to be clinically effective in the stage of the condition you are examining, though a difficulty here is the relative lack of such agents.

7. Conclusions

Animal models have made a significant contribution to our understanding of the changes that occur in neurodegenerative disorders, particularly Alzheimer's disease, as they have been used to determine underlying pathological mechanisms, determine the changes induced by genetic modification and to evaluate the efficacy of putative therapeutic agents. However, their utility is limited as they are not able to recapitulate all aspects of the complex neurodegenerative changes and positive pre-clinical results have not yet translated to the clinic. While one possible reason for this may be a lack of external validity or generalisability of a model, it is also clear that there is also a bias in publication that results in pre-clinical studies showing over-promising effects in the clinic [140,7]. While there may be many reasons for this, one possibility is the fact that it is easier to publish positive results rather than neutral or negative results, this publication bias still reduces their applicability to the clinical situation. Furthermore, the complexity of behavioural models coupled with procedural and interpretational issues have cast doubt on the usefulness of such models for predicting clinical efficacy

Based on the above, and reviews by others, there are a number of key considerations to improve the translation from animal models to the clinic. First, choosing a model, careful design of the experiment and being aware of its limitations is critical. The models can tell us a lot about, for example, mechanistic pathways in a particular neurodegenerative disease, but cannot represent the full complex disorder. Second, having chosen the model, a clearly defined and validated protocol should be followed for the model and careful notice taken of the ARRIVE guidelines. Following clear consistent steps will ensure that better and more realistic pre-clinical studies are conducted and published, which will lead to greater translation to the clinic.

Funding

Declaration of Competing Interest

None.

References

[1] M.J. Prince, World Alzheimer Report 2015: the Global Impact of Dementia (published online Aug 25) https://www.alz.co.uk/research/world-report-2015 Date Accessed: March 29, (2019).

[2] P.F. Durrenberger, F.S. Fernando, S.N. Kashefi, T.P. Bonnert, D. Seilhean, B. Nait-Oumesmar, A. Schmitt, P.J. Gebicke-Haerter, P. Falkai, E. Grünblatt, M. Palkovits, T. Arzberger, H. Kretzschmar, D.T. Dexter, R. Reynolds, Common mechanisms in neurodegeneration and neuroinflammation: a BrainNet Europe gene expression microarray study, J. Neural Transm. 122 (2015) 1055–1068.

[3] D.M. Smith, Could a common mechanism of protein degradation impairment underlie many neurodegenerative diseases? J. Exp. Neurosci. 12 (2018) 1179069518794675, , https://doi.org/10.1177/1179069518794675.

[4] A. Xie, J. Gao, L. Xu, D. Meng, Shared mechanisms of neurodegeneration in Alzheimer's disease and Parkinson's disease, Biomed. Res. Int. 2014 (2014), https://doi.org/10.1155/2014/648740 648740.

[5] H.R. Amanatkar, B. Papagiannopoulos, G.T. Grossberg, Analysis of recent failures of disease modifying therapies in Alzheimer's disease suggesting a new methodology for future studies, Expert Rev. Neurother. 17 (1) (2017) 7–16.

[6] A. Shuaib, K.R. Lees, P. Lyden, J. Grotta, A. Davalos, et al., NXY-059 for the treatment of acute ischemic stroke, N. Engl. J. Med. 357 (2007) 562–571.

[7] K.J. Egan, H.M. Vesterinen, V. Beglopoulos, E.S. Sena, M.R. Macleod, From a mouse: systematic analysis reveals limitations of experiments testing interventions in Alzheimer's disease mouse models, Evid. Preclin. Med. 3 (1) (2016) e00015Epub 2016/08/01. pmid:29214041.

[8] C.G. Begley, J.P. Ioannidis, Reproducibility in science: improving the standard for basic and preclinical research, Circ. Res. 116 (1) (2015) 116–126.

[9] K.K. Tsilidis, O.A. Panagiotou, E.S. Sena, E. Aretouli, E. Evangelou, D.W. Howells, R. Al-Shahi Salman, M.R. Macleod, J.P. Ioannidis, Evaluation of excess significance bias in animal studies of neurological diseases, PLoS Biol. 11 (7) (2013) e1001609.

[10] H.B. Van der Worp, P. de Haan, E. Morrema, C.J. Kalkman, Methodological quality of animal studies on neuroprotection in focal cerebral ischaemia, J. Neurol. 252 (2005) 1108–1114.

[11] B.P. Kirby, Animal models of psychotic disorders: dimensional approach modeling negative symptoms, in: M. Pletnikov, J. Waddington (Eds.), Modelling Psychopathological Dimensions of Schizophrenia, Elsevier, 2016, pp. 55–67.

[12] N.A. Crossley, E. Sena, J. Goehler, J. Horn, B. van der Worp, P.M.W. Bath, M. Macleod, U. Dirnagl, Empirical evidence of bias in the design of experimental stroke studies. A metaepidemiologic approach, Stroke 39 (2008) 929–934.

[13] E. Asan, D.M. Yilmazer-Hanke, M. Eliava, M. Hantsch, K.P. Lesch, A. Schmitt, The corticotropin-releasing factor (CRF)-system and monoaminergic afferents in the central amygdala: investigations in different mouse strains and comparison with the rat, Neuroscience 131 (4) (2005) 953–967.

[14] W.D. Hirst, B. Abrahamsen, F.E. Blaney, A.R. Calver, L. Aloj, G.W. Price, A.D. Medhurst, Differences in the central nervous system distribution and pharmacology of the mouse 5-hydroxytryptamine-6 receptor compared with rat and human receptors investigated by radioligand binding, site-directed mutagenesis, and molecular modeling, Mol. Pharmacol. 64 (2003) 1295–1308.

[15] J. Snyder, A. Soumier, M. Brewer, J. Pickel, H.A. Cameron, Adult-born hippocampal neurons are more numerous, faster maturing, and more involved in behavior in rats than in mice, J. Neurosci. 29 (2009) 14484–14495.

[16] K.M. Frick, E.T. Stillner, J. Berger-Sweeney, Mice are not little rats: species differences in a one-day water maze task, Neuroreport 11 (2000) 3461–3465.

[17] R. Gerlai, N.S. Clayton, Analysing hippocampal function in transgenic mice: an ethological perspective, Trends Neurosci. 22 (1999) 47–51.

[18] C. Solinsky, B.P. Kirby, Medial prefrontal cortex lesions in mice do not impair effort-based decision making, Neuropharmacology 65 (2013) 223–231, https://doi.org/10.1016/j.neuropharm.2012.10.005.

[19] G.W. Bothe, V.J. Bolivar, M.J. Vedder, J.G. Geistfeld, Behavioral differences among fourteen inbred mouse strains commonly used as disease models, Comp. Med. 55 (4) (2005) 326–334.

[20] V. Voikar, S. Kõks, E. Vasar, H. Rauvala, Strain and gender differences in the behavior of mouse lines commonly used in transgenic studies, Physiol. Behav. 72 (1–2) (2001) 271–281.

[21] C.V. Vorhees, M.T. Williams, Reprint of "Value of water mazes for assessing spatial and egocentric learning and memory in rodent basic research and regulatory studies", Neurotoxicol. Teratol. 52 (2015) 93–108.

[22] L.A. Galea, K.A. Uban, J.R. Epp, S. Brummelte, C.K. Barha, W.L. Wilson, et al., Endocrine regulation of cognition and neuroplasticity: our pursuit to unveil the complex interaction between hormones, the brain, and behavior, Can. J. Exp. Psychol. 62 (2008) 247–260.

[23] J.E. Gresack, K.M. Frick, Male mice exhibit better spatial working and reference memory than females in a water-escape radial arm maze task, Brain Res. 982 (2003) 98–107.

[24] L.A. Galea, S.R. Wainwright, M.M. Roes, P. Duarte-Guterman, C. Chow, D.K. Hamson, Sex, hormones, and neurogenesis in the hippocampus: hormonal modulation of neurogenesis and potential functional implications, J. Neuroendocrinol. 25 (2013) 1035–1061.

[25] N. Benaroya-Milshtein, N. Hollander, A. Apter, T. Kukulansky, N. Raz, A. Wilf, I. Yaniv, C.G. Pick, Environmental enrichment in mice decreases anxiety, attenuates stress responses and enhances natural killer cell activity, Eur. J. Neurosci. 20 (2004) 1341–1347.

[26] T.J. Lambert, S.M. Fernandez, K.M. Frick, Different types of environmental enrichment have discrepant effects on spatial memory and synaptophysin levels in female mice, NeurobioL. Learn. Mem. 83 (2005) 206–216.

[27] R. López-Aumatell, E. Martínez-Membrives, E. Vicens-Costa, T. Cañete, G. Blázquez, C. Mont-Cardona, et al., Effects of environmental and physiological covariates on sex differences in unconditioned and conditioned anxiety and fear in a large sample of genetically heterogeneous (N/Nih-HS) rats, Behav. Brain Funct. 7 (2011) 48, https://doi.org/10.1186/1744-9081-7-48.

[28] N. Andrews, S.E. File, Handling history of rats modifies behavioural effects of drugs in the elevated plus-maze test of anxiety, Eur. J. Pharmacol. 235 (1993) 109–112.

[29] J.C. Crabbe, D. Wahlsten, B.C. Dudek, Genetics of mouse behavior: interactions with laboratory environment, Science 284 (1999) 1670–1672.

[30] J. Götz, L.G. Bodea, M. Goedert, Rodent models for Alzheimer disease, Nat. Rev. Neurosci. 19 (10) (2018) 583–598.

[31] World Health Organisation, The Epidemiology and Impact of Dementia, WHO, 2015 (Accessed 29th March 2019), http://www.who.int/mental_health/neurology/dementia/dementia_thematicbrief_epidemiology.pdf.

[32] M. Prince, M. Prina, M. Guerchet, World Alzheimer Report, Journey of Caring. An Analysis of Long-Term Care for Dementia, Alzheimer's Disease International, London, 2013.

[33] American Psychiatric Association, Diagnostic and Statistical Manual of Mental Disorders (DSM-5), (2013) (Accessed 7th March 2019), http://www.psychiatry.org/dsm5.

[34] G. Coughlan, J. Laczó, J. Hort, A.M. Minihane, M. Hornberger, Spatial navigation deficits – overlooked cognitive marker for preclinical Alzheimer disease? Nat. Rev. Neurol. 14 (8) (2018) 496–506.

[35] D.N. Drechsel, A.A. Hyman, M.H. Cobb, M.W. Kirschner, Modulation of the dynamic instability of tubulin assembly by the microtubule-associated protein tau, Mol. Biol. Cell 3 (10) (1992) 1141–1154.

[36] M. Medina, J. Avila, The role of extracellular Tau in the spreading of neurofibrillary pathology, Front. Cell. Neurosci. 8 (2014) 113.

[37] S.A. Frautschy, F. Yang, L. Calderon, G.M. Cole, Rodent models of Alzheimer's disease: rat A beta infusion approaches to amyloid deposits, Neurobiol. Aging 17

(1996) 311–321.

[38] T.P. O'Leary, R.E. Brown, Visuo-spatial learning and memory deficits on the Barnes maze in the 16-month-old APPswe/PS1dE9 mouse model of Alzheimer's disease, Behav. Brain Res. 201 (1) (2009) 120–127.

[39] S.J. Webster, A.D. Bachstetter, P.T. Nelson, F.A. Schmitt, L.J. Van Eldik, Using mice to model Alzheimer's dementia: an overview of the clinical disease and the preclinical behavioral changes in 10 mouse models, Front. Genet. 23 (5) (2014) 88.

[40] T. Urano, C. Tohda, Icariin improves memory impairment in Alzheimer's disease model mice (5xFAD) and attenuates amyloid β-induced neurite atrophy, Phytother. Res. 24 (11) (2010) 1658–1663.

[41] M.A. Westerman, D. Cooper-Blacketer, A. Mariash, et al., The relationship between Abeta and memory in the Tg2576 mouse model of Alzheimer's disease, J. Neurosci. 22 (2002) 1858–1867.

[42] C.C. Wu, F. Chawla, D. Games, et al., Selective vulnerability of dentate granule cells prior to amyloid deposition in PDAPP mice: digital morphometric analyses, Proc. Natl. Acad. Sci. U. S. A. 101 (2004) 7141–7146.

[43] D.L. King, G.W. Arendash, F. Crawford, T. Sterk, J. Menendez, M.J. Mullan, Progressive and gender-dependent cognitive impairment in the APPSW transgenic mouse model for Alzheimer's disease, Behav. Brain Res. 103 (1999) 145–162.

[44] S. Oddo, A. Caccamo, J.D. Shepherd, M.P. Murphy, T.E. Golde, R. Kayed, R. Metherate, M.P. Mattson, Y. Akbari, F.M. Laferla, Triple-transgenic model of Alzheimer's disease with plaques and tangles: intracellular Abeta and synaptic dysfunction, Neuron 39 (2003) 409–421.

[45] R.G.M. Morris, Spatial localization does not require the presence of local cues, Learn. Motiv. 12 (1981) 239–260.

[46] A.S. Murtishaw, C.F. Heaney, M.M. Bolton, K.C.D. Belmonte, M.A. Langhardt, J.W. Kinney, Intermittent streptozotocin administration induces behavioral and pathological features relevant to Alzheimer's disease and vascular dementia, Neuropharmacology 137 (2018) 164–177.

[47] F.R. Farina, T. Burke, D. Coyle, K. Jeter, M. McGee, J. O'Connell, D. Taheny, S. Commins, Learning efficiency: the influence of cue salience during spatial navigation, Behav. Processes 116 (2015) 17–27.

[48] K.L. Shires, J.P. Aggleton, Mapping immediate-early gene activity in the rat after place learning in a water-maze: the importance of matched control conditions, Eur. J. Neurosci. 28 (5) (2008) 982–996.

[49] R.G. Morris, P. Garrud, J.N. Rawlins, J. O'Keefe, Place navigation impaired in rats with hippocampal lesions, Nature 297 (5868) (1982) 681–683.

[50] M. Diviney, D. Fey, S. Commins, Hippocampal contribution to vector model hypothesis during cue-dependent navigation, Learn. Mem. 20 (7) (2013) 367–378.

[51] R. Baeta-Corral, L. Giménez-Llort, Persistent hyperactivity and distinctive strategy features in the Morris water maze in 3xTg-AD mice at advanced stages of disease, Behav. Neurosci. 129 (2) (2015) 129–137.

[52] M. Hick, U. Herrmann, S.W. Weyer, J.P. Mallm, J.A. Tschäpe, M. Borgers, M. Mercken, F.C. Roth, A. Draguhn, L. Slomianka, D.P. Wolfer, M. Korte, U.C. Müller, Acute function of secreted amyloid precursor protein fragment APPsα in synaptic plasticity, Acta Neuropathol. 129 (1) (2015) 21–37.

[53] D.S. Olton, The radial arm maze as a tool in behavioral pharmacology, Physiol. Behav. 40 (6) (1987) 793–797.

[54] S.J. Webster, A.D. Bachstetter, L.J. Van Eldik, Comprehensive behavioral characterization of an APP/PS-1 double knock-in mouse model of Alzheimer's disease, Alzheimers Res. Ther. 5 (3) (2013) 28.

[55] D. Ryan, D. Koss, E. Porcu, H. Woodcock, L. Robinson, B. Platt, G. Riedel, Spatial learning impairments in PLB1Triple knock-in Alzheimer mice are task-specific and age-dependent, Cell. Mol. Life Sci. 70 (14) (2013) 2603–2619.

[56] S.T. Anderson, S. Commins, P.N. Moynagh, A.N. Coogan, Lipopolysaccharide-induced sepsis induces long-lasting affective changes in the mouse, Brain Behav. Immun. 43 (2015) 98–109.

[57] V. Capurro, P. Busquet, J.P. Lopes, R. Bertorelli, G. Tarozzo, M.L. Bolognesi, D. Piomelli, A. Reggiani, A. Cavalli, Pharmacological characterization of memoquin, a multi-target compound for the treatment of Alzheimer's disease, PLoS One 8 (2) (2013) e56870.

[58] N. Suganthy, D.S. Malar, K.P. Devi, Rhizophora mucronata attenuates beta-amyloid induced cognitive dysfunction, oxidative stress and cholinergic deficit in Alzheimer's disease animal model, Metab. Brain Dis. 31 (4) (2016) 937–949.

[59] D.K. Mostafa, C.A. Ismail, D.A. Ghareeb, Differential metformin dose-dependent effects on cognition in rats: role of Akt, Psychopharmacology (Berl.) 233 (13) (2016) 2513–2524.

[60] N. Maroof, S. Ravipati, M.C. Pardon, D.A. Barrett, D.A. Kendall, Reductions in endocannabinoid levels and enhanced coupling of cannabinoid receptors in the striatum are accompanied by cognitive impairments in the AβPPswe/PS1ΔE9 mouse model of Alzheimer's disease, J. Alzheimers Dis. 42 (1) (2014) 227–245.

[61] M.J. Young, P.C. Geiszler, M.C. Pardon, A novel role for the immunophilin FKBP52 in motor coordination, Behav. Brain Res. 313 (2016) 97–110.

[62] D.P. Devanand, X. Liu, P.J. Brown, Impact of functional deficits in instrumental activities of daily living in mild cognitive impairment: a clinical algorithm to predict progression to dementia, Alzheimer Dis. Assoc. Disord. 31 (1) (2017) 55–61.

[63] T.P. O'Leary, A. Robertson, P.H. Chipman, V.F. Rafuse, R.E. Brown, Motor function deficits in the 12 month-old female 5xFAD mouse model of Alzheimer's disease, Behav. Brain Res. 337 (2018) 256–263.

[64] M. Hullmann, C. Albrecht, D. van Berlo, M.E. Gerlofs-Nijland, T. Wahle, A.W. Boots, J. Krutmann, F.R. Cassee, T.A. Bayer, R.P.F. Schins, Diesel engine exhaust accelerates plaque formation in a mouse model of Alzheimer's disease, Part. Fibre Toxicol. 14 (1) (2017) 35.

[65] P.M. Ardestani, A.K. Evans, B. Yi, T. Nguyen, L. Coutellier, M. Shamloo,

Modulation of neuroinflammation and pathology in the 5XFAD mouse model of Alzheimer's disease using a biased and selective beta-1 adrenergic receptor partial agonist, Neuropharmacology 116 (2017) 371–386.

[66] R. Sterniczuk, M.C. Antle, F.M. Laferla, R.H. Dyck, Characterization of the 3xTg-AD mouse model of Alzheimer's disease: part 2. Behavioral and cognitive changes, Brain Res. 1348 (2010) 149–155.

[67] K. Plucińska, B. Crouch, D. Koss, L. Robinson, M. Siebrecht, G. Riedel, B. Platt, Knock-in of human BACE1 cleaves murine APP and reiterates Alzheimer-like phenotypes, J. Neurosci. 34 (32) (2014) 10710–10728.

[68] S. Rahayel, J. Frasnelli, S. Joubert, The effect of Alzheimer's disease and Parkinson's disease on olfaction: a meta-analysis, Behav. Brain Res. 231 (2012) 60–74.

[69] W. Xu, M. Lopez-Guzman, C. Schoen, S. Fitzgerald, S.L. Lauer, R.A. Nixon, E. Levy, D.A. Wilson, Spared piriform cortical single-unit odor processing and odor discrimination in the Tg2576 mouse model of Alzheimer's disease, PLoS One 9 (9) (2014) e106431.

[70] Z. Li, K. Jia, Y. Duan, D. Wang, Z. Zhou, S. Dong, Xanomeline derivative EUK1001 attenuates Alzheimer's disease pathology in a triple transgenic mouse model, Mol. Med. Rep. 16 (5) (2017) 7835–7840.

[71] F.T. Meng, J. Zhao, H. Fang, L.F. Zhang, H.M. Wu, Y.J. Liu, Upregulation of mineralocorticoid receptor in the hypothalamus associated with a high anxiety-like level in apolipoprotein E4 transgenic mice, Behav. Genet. 47 (4) (2017) 416–424.

[72] H.L. Wang, X.H. Xian, Q.Y. Song, C. Pang, J.L. Wang, M.W. Wang, W.B. Li, Age-related alterations of neuronal excitability and voltage-dependent Ca2+ current in a spontaneous mouse model of Alzheimer's disease, Behav. Brain Res. 321 (2017) 209–213.

[73] J. Kassa, J. Bajgar, K. Kuca, D. Jun, Ramesh C. Gupta (Ed.), Behavioral Toxicity of Nerve Agents in Handbook of Toxicology of Chemical Warfare Agents, second edition, Academic Press, London, 2015.

[74] V. Paban, C. Manrique, M. Filali, S. Maunoir-Regimbal, F. Fauvelle, B. Alescio-Lautier, Therapeutic and preventive effects of methylene blue on Alzheimer's disease pathology in a transgenic mouse model, Neuropharmacology 76 (Pt A) (2014) 68–79.

[75] Y. Senechal, P.H. Kelly, K.K. Dev, Amyloid precursor protein knockout mice show age-dependent deficits in passive avoidance learning, Behav. Brain Res. 186 (2008) 126–132.

[76] D.N. Barry, A.N. Coogan, S. Commins, The time course of systems consolidation of spatial memory from recent to remote retention: a comparison of the Immediate Early Genes Zif268, c-Fos and Arc, Neurobiol. Learn. Mem. 128 (2016) 46–55.

[77] R.M.J. Deacon, A novel approach to discovering treatments for Alzheimer's disease, Alzheimers Dis. Parkinsonism 4 (2) (2014), https://doi.org/10.4172/2161-0460.1000142 2014.

[78] J. O'Keefe, J. Dostrovsky, The hippocampus as a spatial map. Preliminary evidence from unit activity in the freely-moving rat, Brain Res. 34 (1) (1971) 171–175.

[79] T. Hafting, M. Fyhn, S. Molden, M.B. Moser, E.I. Moser, Microstructure of a spatial map in the entorhinal cortex, Nature 436 (7052) (2005) 801–806.

[80] H. Braak, E. Braak, Staging of Alzheimer's disease-related neurofibrillary changes, Neurobiol. Aging 16 (3) (1995) 271–278.

[81] M. Krause, Z. Yang, G. Rao, F.P. Houston, C.A. Barnes, Altered dendritic integration in hippocampal granule cells of spatial learning-impaired aged rats, J. Neurophysiol. 99 (6) (2008) 2769–2778.

[82] M. De Coninck, D. Van Dam, C. Van Ginneken, P.P. De Deyn, Adapted Morris Water Maze protocol to prevent interference from confounding motor deficits on cognitive functioning, Somatosens. Mot. Res. 34 (3) (2017) 172–178.

[83] A.M. Gouirand, L. Matuszewich, The effects of chronic unpredictable stress on male rats in the water maze, Physiol. Behav. 86 (1–2) (2005) 21–31.

[84] R.L. Spencer, W.K. O'Steen, B.S. McEwen, Water maze performance of aged Sprague-Dawley rats in relation to retinal morphologic measures, Behav. Brain Res. 68 (1995) 139–150.

[85] G.C. Tombaugh, W.B. Rowe, G.M. Rose, The slow after hyperpolarization in hippocampal CA1 neurons covaries with spatial learning ability in aged Fisher 344 rats, J. Neurosci. 25 (2005) 2609–2616.

[86] G.W. Arendash, J. Lewis, R.E. Leighty, E. McGowan, J.R. Cracchiolo, M. Hutton, M.F. Garcia, Multi-metric behavioral comparison of APPsw and P301L models for Alzheimer's disease: linkage of poorer cognitive performance to tau pathology in forebrain, Brain Res. 1012 (1-2) (2004) 29–41.

[87] D. Moechars, I. Dewachter, K. Lorent, D. Reversé, V. Baekelandt, A. Naidu, I. Tesseur, K. Spittaels, C.V. Haute, F. Checler, E. Godaux, B. Cordell, F. Van Leuven, Early phenotypic changes in transgenic mice that overexpress different mutants of amyloid precursor protein in brain, J. Biol. Chem. 274 (10) (1999) 6483–6492.

[88] R. Lalonde, K. Fukuchi, C. Strazielle, Neurologic and motor dysfunctions in APP transgenic mice, Rev. Neurosci. 23 (4) (2012) 363–379.

[89] D.N. Barry, S. Commins, A novel control condition for spatial learning in the Morris water maze, J. Neurosci. Methods 318 (2019) 1–5.

[90] J. Kealy, S. Commins, The rat perirhinal cortex: a review of anatomy, physiology, plasticity, and function, Prog. Neurobiol. 93 (4) (2011) 522–548.

[91] S.J. Cohen, R.W.Jr. Stackman, Assessing rodent hippocampal involvement in the novel object recognition task. A review, Behav. Brain Res. 285 (2015) 105–117.

[92] S.J. Holmes, F. Jane Fitch, A.W. Ellis, Age of acquisition affects object recognition and naming in patients with Alzheimer's disease, J. Clin. Exp. Neuropsychol. 28 (6) (2006) 1010–1022.

[93] S. Laatu, A. Revonsuo, H. Jäykkä, R. Portin, J.O. Rinne, Visual object recognition in early Alzheimer's disease: deficits in semantic processing, Acta Neurol. Scand. 108 (2) (2003) 82–89.

[94] T.C.C. Pinto, L. Machado, T.M. Bulgacov, A.L. Rodrigues-Júnior, M.L.G. Costa,

R.C.C. Ximenes, E.B. Sougey, Is the Montreal Cognitive Assessment (MoCA) screening superior to the Mini-Mental State Examination (MMSE) in the detection of mild cognitive impairment (MCI) and Alzheimer's Disease (AD) in the elderly? Int. Psychogeriatr. 14 (2018) 1–14.

[95] D.C. Delis, J.H. Kramer, E. Kaplan, B.A. Ober, CVLT, California Verbal Learning Test: Adult Version: Manual, Psychological Corporation, 1987.

[96] O. Bezdicek, D.J. Libon, H. Stepankova, E. Panenkova, J. Lukavsky, K.D. Garrett, et al., Development, validity, and normative data study for the 12-word Philadelphia Verbal Learning Test [czP(r)VLT-12] among older and very old Czech adults, Clin. Neuropsychol. 28 (7) (2014) 1162–1181.

[97] U. Díaz-Orueta, A. Blanco-Campal, T. Burke, Rapid review of cognitive screening instruments in MCI: proposal for a process-based approach modification of over-lapping tasks in select widely used instruments, Int. Psychogeriatr. 30 (5) (2018) 663–672.

[98] J.S. Snowden, J.C. Thompson, D. Neary, Knowledge of famous faces and names in semantic dementia, Brain 127 (Pt 4) (2004) 860–872.

[99] R.E. Curiel, E. Crocco, M. Rosado, R. Duara, M.T. Greig, A. Raffo, D.A. Loewenstein, A brief computerized paired associate test for the detection of mild cognitive impairment in community-dwelling older adults, J. Alzheimers Dis. 54 (Sep (2)) (2016) 793–799 2016.

[100] S. Hsieh, S. Schubert, C. Hoon, E. Mioshi, J.R. Hodges, Validation of the Addenbrooke's cognitive examination III in frontotemporal dementia and Alzheimer's disease, Dement. Geriatr. Cogn. Disord. 36 (3–4) (2013) 242–250.

[101] Z.S. Nasreddine, N.A. Phillips, V. Bédirian, S. Charbonneau, V. Whitehead, I. Collin, J.L. Cummings, H. Chertkow, The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment, J. Am. Geriatr. Soc. 53 (4) (2005) 695–699.

[102] R.W. Hopkins, L.A. Kilik, D.J. Day, C.P. Rows, P.F. Hamilton, The brief Kingston standardized cognitive assessment–revised, Int. J. Geriatr. Psychiatry 20 (3) (2005) 227–231.

[103] E. Kokmen, J.M. Naessens, K.P. Offord, A short test of mental status: description and preliminary results, Mayo Clin. Proc. 62 (4) (1987) 281–288.

[104] B. Dubois, H.H. Feldman, C. Jacova, H. Hampel, J.L. Molinuevo, K. Blennow, S.T. DeKosky, S. Gauthier, D. Selkoe, R. Bateman, S. Cappa, S. Crutch, S. Engelborghs, G.B. Frisoni, N.C. Fox, D. Galasko, M.O. Habert, G.A. Jicha, A. Nordberg, F. Pasquier, G. Rabinovici, P. Robert, C. Rowe, S. Salloway, M. Sarazin, S. Epelbaum, L.C. de Souza, B. Vellas, P.J. Visser, L. Schneider, Y. Stern, P. Scheltens, J.L. Cummings, Advancing research diagnostic criteria for Alzheimer's disease: the IWG-2 criteria, Lancet Neurol. 13 (6) (2014) 614–629.

[105] M. Irish, O. Piguet, J.R. Hodges, M. Hornberger, Common and unique gray matter correlates of episodic memory dysfunction in frontotemporal dementia and Alzheimer's disease, Hum. Brain Mapp. 35 (4) (2014) 1422–1435.

[106] M.J. Eacott, G. Norman, Integrated memory for object, place, and context in rats: a possible model of episodic-like memory? J. Neurosci. 24 (2004) 1948–1953.

[107] N.S. Clayton, A. Dickinson, Motivational control of caching behaviour in the scrub jay, Aphelocoma coerulescens, Anim. Behav. 57 (1999) 435–444.

[108] E. Dere, J.P. Huston, M.A. De Souza Silva, Episodic-like memory in mice: simultaneous assessment of object, place and temporal order memory, Brain Res. Brain Res. Protoc. 16 (1–3) (2005) 10–19.

[109] M.A. Good, G. Hale, V. Staal, Impaired "episodic-like" object memory in adult APPswe transgenic mice, Behav. Neurosci. 121 (2) (2007) 443–448.

[110] K.E. Davis, A. Easton, M.J. Eacott, J. Gigg, Episodic-like memory for what-where-which occasion is selectively impaired in the 3xTgAD mouse model of Alzheimer's disease, J. Alzheimers Dis. 33 (3) (2013) 681–698.

[111] M. Carli, T.W. Robbins, J.L. Evenden, B.J. Everitt, Effects of lesions to ascending noradrenergic neurones on performance of a 5-choice serial reaction task in rats; implications for theories of dorsal noradrenergic bundle function based on selective attention and arousal, Behav. Brain Res. 9 (3) (1983) 361–380.

[112] L.H. Beck, E.D. Bransome Jr., A.F. Mirsky, H.E. Rosvold, I. Sarason, A continuous performance test of brain damage, J. Consult. Psychol. 20 (5) (1956) 343–350.

[113] I.H. Robertson, T. Manly, J. Andrade, B.T. Baddeley, J. Yiend, 'Oops!': performance correlates of everyday attentional failures in traumatic brain injured and normal subjects, Neuropsychologia 35 (6) (1997) 747–758.

[114] C. Romberg, M.P. Mattson, M.R. Mughal, T.J. Bussey, L.M. Saksida, Impaired attention in the 3xTgAD mouse model of Alzheimer's disease: rescue by donepezil (Aricept), J. Neurosci. 31 (9) (2011) 3500–3507.

[115] A.V. Bharmal, B.A. Kent, T.J. Bussey, L.M. Saksida, Performance of transgenic TgTau-P301L mice in a 5-choice serial reaction time task (5-CSRTT) as a model of Alzheimer's disease, Psychiatr. Danub. 27 (Suppl. 1) (2015) S515–25.

[116] C. van Doorn, A.L. Gruber-Baldini, S. Zimmerman, et al., Dementia as a risk factor for falls and fall injuries among nursing home residents, J. Am. Geriatr. Soc. 51 (2003) 1213–1218.

[117] K.R. Stover, M.A. Campbell, C.M. Van Winssen, R.E. Brown, Analysis of motor function in 6-month-old male and female 3xTg-AD mice, Behav. Brain Res. 281 (2015) 16–23.

[118] J.M. Hausdorff, I. Hillel, S. Shustak, S. Del Din, E.M.J. Bekkers, E. Pelosin, F. Nieuwhof, L. Rochester, A. Mirelman, Everyday stepping quantity and quality among older adult fallers with and without mild cognitive impairment: initial evidence for new motor markers of cognitive deficits? J. Gerontol. A Biol. Sci. Med. Sci. 73 (8) (2018) 1078–1082.

[119] M. Montero-Odasso, F. Pieruccini-Faria, et al., Motor phenotype in neurodegenerative disorders: gait and balance platform study design protocol for the ontario neurodegenerative research initiative (ONDRI), J. Alzheimers Dis. 59 (2) (2017) 707–721.

[120] E.A. Walshe, M.R. Patterson, S. Commins, R.A. Roche, Dual-task and electrophysiological markers of executive cognitive processing in older adult gait and

fall-risk, Front. Hum. Neurosci. 9 (2015) 200.

[121] S.W. Muir, M. Speechley, J. Wells, M. Borrie, K. Gopaul, M. Montero-Odasso, Gait assessment in mild cognitive impairment and Alzheimer's disease: the effect of dual-task challenges across the cognitive spectrum, Gait Posture 35 (1) (2012) 96–100.

[122] A. Videnovic, A.S. Lazar, R.A. Barker, S. Overeem, 'The clocks that time us'–circadian rhythms in neurodegenerative disorders, Nat. Rev. Neurol. 10 (2014) 683–693.

[123] Y. Leng, E.S. Musiek, K. Hu, F.P. Cappuccio, K. Yaffe, Association between circadian rhythms and neurodegenerative diseases, Lancet Neurol. 18 (3) (2019) 307–318.

[124] S.L. Naismith, I.B. Hickie, Z. Terpening, et al., Circadian misalignment and sleep disruption in mild cognitive impairment, J. Alzheimers Dis. 38 (2014) 857–866.

[125] N. Cermakian, E.W. Lamont, P. Boudreau, D.B. Boivin, Circadian clock gene expression in brain regions of Alzheimer's disease patients and control subjects, J. Biol. Rhythms 26 (2011) 160–170.

[126] B.S. Diniz, M.A. Butters, S.M. Albert, M.A. Dew, C.F. Reynolds 3rd, Late-life depression and risk of vascular dementia and Alzheimer's disease: systematic review and meta-analysis of community-based cohort studies, Br. J. Psychiatry 202 (5) (2013) 329–335.

[127] R.D. Porsolt, G. Brossard, C. Hautbois, S. Roux, Rodent models of depression: forced swimming and tail suspension behavioral despair tests in rats and mice, Curr. Protoc. Neurosci. (2001) Chapter 8:Unit 8.10A.

[128] S.T. Anderson, S. Commins, P. Moynagh, A.N. Coogan, Chronic fluoxetine treatment attenuates post-septic affective changes in the mouse, Behav. Brain Res. 297 (2016) 112–115.

[129] D.M. Iascone, S. Padidam, M.S. Pyfer, X. Zhang, L. Zhao, J. Chin, Impairments in neurogenesis are not tightly linked to depressive behavior in a transgenic mouse model of Alzheimer's disease, PLoS One 8 (11) (2013) e79651.

[130] Y. D'Souza, A. Elharram, R. Soon-Shiong, R.D. Andrew, B.M. Bennett, Characterization of Aldh2 (-/-) mice as an age-related model of cognitive impairment and Alzheimer's disease, Mol. Brain 8 (27) (2015).

[131] C.F. Mactutus, R.M. Booze, Accuracy of spatial navigation: the role of platform and tank size, Soc. Neurosci. Abst. 20 (1994) 1014.

[132] A.M. Sabogal-Guáqueta, J.I. Muñoz-Manco, J.R. Ramírez-Pineda, M. Lamprea-Rodriguez, E. Osorio, G.P. Cardona-Gómez, The flavonoid quercetin ameliorates Alzheimer's disease pathology and protects cognitive and emotional function in aged triple transgenic Alzheimer's disease model mice, Neuropharmacology 93 (2015) 134–145.

[133] J.T. Rick, M.P. Murphy, G.O. Ivy, N.W. Milgram, Short intertrial intervals impair water maze performance in old Fischer 344 rats, J. Gerontol. A Biol. Sci. Med. Sci. 51 (4) (1996) B253–60.

[134] H.1 Iivonen, L. Nurminen, M. Harri, H. Tanila, J. Puoliväli, Hypothermia in mice tested in Morris water maze, Behav. Brain Res. 141 (2) (2003) 207–213.

[135] K.E. Davis, K. Burnett, J. Gigg, Water and T-maze protocols are equally efficient methods to assess spatial memory in 3xTg Alzheimer's disease mice, Behav. Brain Res. 331 (2017) 54–66.

[136] A. Olmos-Alonso, S.T. Schetters, S. Sri, K. Askew, R. Mancuso, M. Vargas-Caballero, C. Holscher, V.H. Perry, D. Gomez-Nicola, Pharmacological targeting of CSF1R inhibits microglial proliferation and prevents the progression of Alzheimer's-like pathology, Brain 139 (Pt 3) (2016) 891–907.

[137] Carol Kilkenny, William J. Browne, Innes C. Cuthill, Michael Emerson, Douglas G. Altman, The ARRIVE guidelines: animal research: reporting in vivo experiments, PLoS Biol. 8 (6) (2010) e1000412.

[138] H.M. Vesterinen, E.S. Sena, C. ffrench-Constant, A. Williams, S. Chandran, M.R. Macleod, Improving the translational hit of experimental treatments in multiple sclerosis, Mult. Scler. 16 (2010) 1044–1055.

[139] R. Rosenthal, Experimenter Effects in Behavioral Research, Appleton-Century-Crofts, East Norwalk, CT, US, 1966.

[140] D.A. Korevaar, L. Hooft, G. ter Riet, Systematic reviews and meta-analyses of preclinical studies: publication bias in laboratory animal experiments, Lab. Anim. 45 (2011) 225–230.

[141] L.M. Luh, I. Das, A. Bertolotti, qMotor, a set of rules for sensitive, robust and quantitative measurement of motor performance in mice, Nat. Protoc. 12 (7) (2017) 1451–1457.

[142] R.M.J. Deacon, Measuring motor coordination in mice, J. Vis. Exp. 75 (2013) e2609.

[143] G.M. Rozas, J. Guerra, J.L. Labandeira-García, An automated rotarod method for quantitative drug-free evaluation of overall motor deficits in rat models of parkinsonism, Brain Res. Protoc. 2 (1) (1997) 75–84.

[144] T.N. Luong, H.J. Carlisle, A. Southwell, P.H. Patterson, Assessment of motor balance and coordination in mice using the balance beam, J. Vis. Exp. 49 (2376) (2011).

[145] R.J. Carter, A.J. Morton, S.B. Dunnett, Motor coordination and balance in rodents, Curr. Protoc. Neurosci. 15 (2001) 8.12.1–8.12.14, https://doi.org/10.1002/0471142301.ns0812s15.

[146] C.V. Vorhees, M.T. Williams, Morris water maze: procedures for assessing spatial and related forms of learning and memory, Nat. Protoc. 1 (2006) 848–858.

[147] R.M.J. Deacon, N.J. Rawlins, T-maze alternation in the rodent, Nat. Protoc. 1 (2006) 7–12.

[148] M. Leger, A. Quiedeville, V. Bouet, B. Haelewyn, M. Boulouard, P. Schumann-Bard, T. Freret, Object recognition test in mice, Nat. Protoc. 8 (12) (2013) 2531–2537.

[149] A.A. Walf, C.A. Frye, The use of the elevated plus maze as an assay of anxiety-related behavior in rodent, Nat. Protoc. 2 (2007) 322–328.