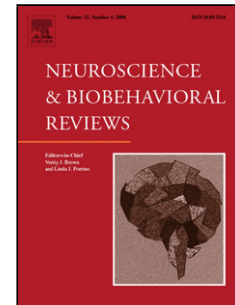


Accepted Manuscript

Title: Evaluating rodent motor functions: Which tests to choose?

Authors: Lisa-Maria Schönfeld, Dearbhaile Dooley, Ali Jahanshahi, Yasin Temel, Sven Hendrix



PII: S0149-7634(17)30441-4
DOI: <https://doi.org/10.1016/j.neubiorev.2017.10.021>
Reference: NBR 2982

To appear in:

Received date: 22-6-2017
Revised date: 18-9-2017
Accepted date: 23-10-2017

Please cite this article as: Schönfeld, Lisa-Maria, Dooley, Dearbhaile, Jahanshahi, Ali, Temel, Yasin, Hendrix, Sven, Evaluating rodent motor functions: Which tests to choose?. Neuroscience and Biobehavioral Reviews <https://doi.org/10.1016/j.neubiorev.2017.10.021>

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Evaluating rodent motor functions:

Which tests to choose?

Lisa-Maria Schönfeld^{a,b1}, Dearbhaile Dooley^{a,d2}, Ali Jahanshahi^b, Yasin

Temel^{b,c#} & Sven Hendrix^{a#}

^a Department of Morphology, Biomedical Research Institute (BIOMED), Hasselt University, Hasselt, Belgium

^b Department of Neuroscience, Maastricht University, Maastricht, the Netherlands

^c Department of Neurosurgery, Maastricht University Medical Center, Maastricht, the Netherlands

^d Laboratory of Experimental Hematology, Vaccine and Infectious Disease Institute (Vaxinfecio), University of Antwerp, Antwerp, Belgium

[#] both authors contributed equally

Corresponding author:

Sven Hendrix

Biomedical Research Institute, Hasselt University

Martelarenlaan 42, 3500 Hasselt, Belgium

sven.hendrix@uhasselt.be

Highlights:

- For high quality results, sensitive behavioral tests for rodents should be used
- Three major categories of rodent motor skill tests are reviewed and evaluated

¹ Present address: Comparative Psychology, Institute of Experimental Psychology, Heinrich-Heine-University Düsseldorf, Düsseldorf, Germany (Email: lisa-maria.schoenfeld@hhu.de)

² Present address: Clinical Medicine, School of Medicine, Trinity College Dublin, Dublin, Ireland

- Recommendations for choosing the best behavioral test are provided

ABSTRACT

Damage to the motor cortex induced by stroke or traumatic brain injury (TBI) can result in chronic motor deficits. For the development and improvement of therapies, animal models which possess symptoms comparable to the clinical population are used. However, the use of experimental animals raises valid ethical and methodological concerns. To decrease discomfort by experimental procedures and to increase the quality of results, non-invasive and sensitive rodent motor tests are needed.

A broad variety of rodent motor tests are available to determine deficits after stroke or TBI. The current review describes and evaluates motor tests that fall into three categories: Tests to evaluate fine motor skills and grip strength, tests for gait and inter-limb coordination and neurological deficit scores. In this review, we share our thoughts on standardized data presentation to increase data comparability between studies. We also critically evaluate current methods and provide recommendations for choosing the best behavioral test for a new research line.

Keywords: Motor tests, rodents, stroke, TBI

1. INTRODUCTION

Insults to the central nervous system, caused by stroke or traumatic brain injury (TBI), can lead to severe motor impairments in gait or fine motor skills. To develop therapies for motor dysfunction, animal models which have a disease pathophysiology comparable to the clinical population are used. The use of non-human primates increases the translational value of the results, since they are physiologically more similar to humans. However, the vast similarity between non-human primates and humans often sparks an ethical discussion, arguing that one should make use of alternative model organisms, which are less similar to humans. This sensitive topic, if and how an ethical or moral status can be assigned to different kinds of mammals, has been reviewed elsewhere (Prescott 2010). To circumvent a number of these ethical concerns, scientists choose smaller animals such as rodents. Additionally, rodents are an attractive choice due to their high reproduction rate and relatively low costs.

1.1 Experimental design according to the 3Rs

Animal research must be performed according to the '3Rs': replacement, reduction and refinement (Russel and Burch 1959). Considerable effort has been invested to try and identify alternative techniques, thereby successfully replacing animal testing (Giri and Bader 2015; Kim et al. 2015a). However, scientific questions about changes in locomotion or other forms of behavior can – up to now - only be addressed in clinical studies or through the use of animal models. Consequently, animal experiments should be continuously refined by choosing and improving behavioral testing methods that are both sensitive and well standardized. Thereby, unnecessary discomfort or even the number of

animal experiments needed, can be reduced in accordance with the 3Rs (Tannenbaum and Bennett 2015).

1.2 Measuring motor functions in experimental animals

Despite overt differences between rodents and primates, both possess a surprisingly similar functional basis of motor behavior (Cenci et al. 2002). Nonetheless, the physiological manifestation of a neurological deficit cannot be translated directly from humans to animals and *vice versa* (Cenci et al. 2002). Rodents and humans differ greatly by means of limb utilization (walking on four paws) and the use different muscle groups. Furthermore, several compensatory mechanisms can take over in the case of damage to a motor pathway (Carmichael et al. 2001; Whishaw 2000). Therefore, highly specific tests are needed to detect the presence of motor impairments in rodents, including impairments not visible in freely moving animals.

The tests discussed in the next sections are used to measure a variety of motor behaviors and have a clear clinical relevance. *Fine motor skills* are vital to everyday functioning and range from brushing our teeth in the morning to setting the alarm clock at night. In addition, disturbed *gait and limb coordination* severely compromise a patient's independence and can lead to a reliance on various types of walking aids.

2. TESTS TO MEASURE FINE MOTOR SKILLS

For humans, fine motor skills, such as picking up a pen or closing a zipper, are vital in daily life. At a first glance, these skills appear difficult to replicate in animal models, given that animals do not perform comparable tasks. Even so,

rodents possess a great capacity to grasp and manipulate small objects. As fine motor skills in rodents are not necessarily observed in the homecage, a challenging testing environment is needed.

In the next section, the following tests will be reviewed:

- Single pellet reaching test: reaching for pellets through a small opening
- Montoya staircase test: reaching for pellets placed on a staircase
- Adhesive removal test: contacting and removing small adhesives underneath the front paws
- Vermicelli handling test: handling and eating of small pasta pieces
- Grid test and ladder rung walking task: walking over an partially open surface
- Cylinder test: exploration of a vertical wall using the forelimbs
- Grip strength test: holding on to a bar in the presence of a pulling force

2.1 Single pellet reaching test

The ‘single pellet reaching test’, or ‘skilled reaching test’, was developed to provide a view on reaching activity in rats and was later adapted to also allow testing of mice (Lai et al. 2015; Whishaw 2000). Animals have to reach for and grasp individual food pellets for a detailed analysis of forelimb motion and potential deviations in pellet eating. The single pellet reaching test has been used to detect long-term forelimb deficits in rat and mouse models of ischemic stroke or TBI (Adkins et al. 2015; Gilmour et al. 2004; Liu et al. 2014; Metz et al. 2005; Nakagawa et al. 2013; Vaagenes et al. 2015; Whishaw 2000). After TBI, rats showed a decreased amount of successful pellet grasping together

with abnormalities in the quality of the reaching action (Adkins et al. 2015), the latter also being found in mice after stroke (Liu et al. 2014).

2.1.1 Testing environment

Animals are placed in a Plexiglas box mounted on a pedestal. Inside that box, a corridor of two Plexiglas walls leads to a slit in the front wall, through which the animals need to reach for food pellets located on a shelf outside. More specifically, the food pellets are placed in small indentations on the left and right edge of the shelf to encourage reaching with the opposite forelimb. Pellet reaching is filmed from the front, bottom and/or side to quantify forelimb motion frame by frame.

2.1.2 Testing and analysis of performance

2.1.2.1 Standard testing procedure

Animals are food deprived to reach 90% of their free feeding weight prior to training and testing in order to increase their motivation to reach for the pellets. During the first week of training, rats learn to reach for pellets, either directly in the reaching chambers or in a slightly modified training environment, the so-called 'food tray task' (Whishaw 2000). In the food tray task, animals are transferred to boxes with metal bars in the front and can reach through the bars into a bigger tray containing food. As soon as animals learn how to reach for pellets through the bars, training is continued in the testing box. In the testing box, food pellets are placed contralaterally to the preferred limb and are replaced either when the pellet was grasped or knocked off the platform. During training, each animal receives 20 pellets per day. After successful pellet

retrieval, a pellet is dropped in the back of the box to distract rats and allow them reposition themselves for the next reach (Whishaw 2000). Testing takes place by again presenting 20 pellets, sometimes preceded by five practice pellets per animal (Starkey et al. 2012).

Grasping performance can be rated by observing the animal during the testing session; however, the gold standard is filming the animal from several angles, to allow offline analysis. Different terminology exists to evaluate the grasping performance of an animal. Originally a 'reach' was defined as an advancement of the paw through the slit. If a rat then grasps a pellet and eats it from the paw, the behavior is scored as a 'hit' (Whishaw 2000). However, when the pellet is knocked off the platform, the reach has failed. The number of hits can be expressed as a percentage of the total number of reaches (Adkins et al. 2015; Metz et al. 2005; Vaagenes et al. 2015). Alternatively, a score of '1' can be assigned when the animal reaches through the slit and eats the pellet, even if grasping the pellet requires several attempts. In the case where the pellet is grasped but dropped inside the box, a score of '0.5' is given, whereas a score of '0' is given when the pellet is knocked off the platform (Starkey et al. 2012). In addition, a success rate can be calculated as the total numeric score divided by the number of reaches (Starkey et al. 2012). One can consider excluding animals that do not fulfill a baseline success rate of 60% or higher (Starkey et al. 2012; Vaagenes et al. 2015) to reduce variation within the groups caused by inter-individual differences in grasping performance.

2.1.2.2 Qualitative assessment

Changes in grasping behavior and body posture can be assessed with different qualitative rating scales (Adkins et al. 2015; Gilmour et al. 2004; Metz et al. 2005; Whishaw 2000). Pellet retrieval consists of a series of single movement components. Through a qualitative analysis, deviant movement components can be identified. Recently, a semi-automated method to analyze reaching movements was proposed, which could detect behavioral deficits after stroke in mice without the need for a time consuming frame-by-frame inspection (Lai et al. 2015).

2.1.3 Pro and Con

By using the single pellet reaching test, the experimenter can perform a detailed analysis of reaching and grasping behavior for each forelimb separately. However, this extensive and subjective analysis may in some cases be considered disadvantageous. Analysis has to occur for each reaching movement separately and at least two blinded observers have to score the animals to minimize subjective influences. In addition, setup dimensions might not allow testing of multiple animals in parallel, which further increases the time needed for test trials. An automated test version, which evaluates reaching performance based on an algorithm (Lai et al. 2015), is being developed to reduce the time needed for analyzing the reaching patterns and reducing the influence of the observer on the dataset. Unfortunately, automated systems are expensive and therefore not widely distributed across labs, which makes data comparison difficult.

Different terminology further complicates the comparison of behavioral outcomes. For example, the term 'reach' can be interpreted as any movement

from sticking the paw through the slit opening, to successful pellet retrieval. Lastly, repeated movement of an impaired forelimb can result in a training effect and improve limb function over time and this may be seen as an unwanted side effect. On the contrary, functional improvement through repeated limb use, or in other words 'physical rehabilitation therapy', has a high clinical relevance, as physical rehabilitation therapy is the method of choice for treating motor impairments in humans.

2.2 Montoya staircase test

The 'Montoya staircase test' was developed as a method to objectively investigate fine motor skills such as reaching and grasping of food pellets (Montoya et al. 1991). It provides a quantitative readout of the individual forelimbs, which requires neither restraint of the contralateral limb nor time-intensive scoring of each reaching movement by an observer (Montoya et al. 1991). In contrast to the single pellet reaching test, performance of both forelimbs using the Montoya staircase test can be measured during the same test session. While the size of the initial staircase box was tailored to adult rats, a smaller version of the testing boxes is also available for mice (Baird et al. 2001).

The Montoya staircase test has been widely used to assess unilateral forelimb deficits in several animal models of cortical damage, such as ischemic or hemorrhagic stroke, unilateral 6-OHDA lesions, TBI and peripheral nerve lesions (Bouet et al. 2007; Harris et al. 2010; Kloth et al. 2006; MacLellan et al. 2013; Pagnussat Ade et al. 2009). After rats received TBI, the number of eaten pellets was significantly decreased compared to control animals. Even when

pellets were successfully grasped and consumed, single components of the reaching behavior, such as eating the pellet from the paw, remained deviant after the lesion (Schönfeld et al. 2016).

2.2.1 Testing environment

The original apparatus consists of a Plexiglas box with a platform located in the rear half of the box. On the left and the right hand side of the platform, a staircase is located with indentations on each step to hold the food pellets (Figure 1). The increasing depth of the staircase makes it more difficult to reach for the lower steps.

2.2.2 Testing and analysis of performance

Animals can be food deprived up to 90% of their free-feeding weight to increase motivation to reach for the food pellets. Although Montoya et al. (1991) reported immediate reaching behavior of rats during the first habituation sessions in the box, most studies include a training period for the animals in order to achieve a steady baseline reaching performance. Animal training lasts between 10 to 21 days and preferably takes place twice daily (Santos et al. 2013; Trueman et al. 2011; Windle et al. 2006). The training sessions can resemble regular testing conditions, but alternatively a more refined habituation approach might be chosen, where the time to expose the animals to the closed box is gradually prolonged (Schönfeld et al. 2016) and the number of pellets per step is increased initially to attract the animal's attention to the steps (Abrous and Dunnett 1994). The rats need to obtain more than 55% of the presented pellets with their preferred limb before reaching performance can be reliably measured

(Pagnussat Ade et al. 2009; Santos et al. 2013; Windle et al. 2006). In the test sessions, either the reaching performance during the last days of training (Abrous and Dunnett 1994; Bouet et al. 2007) or during two separate testing days is averaged (Clarke et al. 2007; MacLellan et al. 2013; Santos et al. 2013; Windle et al. 2006).

2.2.2.1 Standard testing procedure

Both sides are baited with two or three food pellets per stair, followed by placing the animal in the box for 15 minutes and counting the number of remaining pellets after the trial is finished. The side where the animals consume the most pellets can be designated as the preferred side.

2.2.2.2 Speed test

Both sides are baited and reaching performance is assessed during the first 15 seconds to measure the initial preference for one paw over the other (Montoya et al. 1991).

2.2.2.3 Side test or forced paw use

Both sides are baited separately (e.g. in the first test session pellets are placed only on the left staircase and *vice versa* in the second test session) and the animals remain in the testing box for 5 minutes. Using this paradigm, it can be made clear whether decreased pellet consumption on one side is due to a lack of attention or a unilateral motor disability (Abrous and Dunnett 1994; Montoya et al. 1991).

2.2.2.4 Color-coding of pellets

To increase the number of outcome parameters, sucrose pellets can be colored for each step using odorless food coloring. Based on the color, the

experimenter can identify whether the pellets on a particular step were left untouched or had been knocked off from a higher level. The following parameters can be assessed: (1) Number of 'remaining pellets' for pellets still in place at their original location, (2) number of 'eaten pellets', for the number of pellets of a certain color, no longer present at the end of a testing session, (3) number of 'pellets misplaced' for the pellets, which are found on a stair other than their original location and (4) number of 'pellets lost', for pellets displaced in a way making them impossible to retrieve, which is usually on the floor in the back of the testing box (Beaudin et al. 2013; Kloth et al. 2006; Santos et al. 2013).

2.2.2.5 Qualitative motor assessment

Pellet retrieval requires the ability to extend the forelimb, spread the digits, hold the pellet and bring the pellet to the mouth. To unravel deficits in one part of a full movement sequence, a qualitative evaluation of the reaching actions can be performed. During a standard test session, the animals are filmed and reaching actions can be scored offline, where the absence or presence of different behaviors within a typical reach is being scored (Clarke et al. 2007).

2.2.3 Pro and Con

With the staircase test, fine motor skills of both forelimbs can be investigated in one testing session without the need to restrain the contralateral limb. The resulting scores are quantitative readouts of forelimb function that can be quickly obtained. Due to the small size of the staircase boxes and the fact that a standard functional assessment does not require observation of the animals during the sessions, several animals can be tested in parallel. Different testing

paradigms are available to increase sensitivity and the number of readout parameters. Also postural deficits do not interfere with reaching ability, since lying on the platform stabilizes the trunk of the animals (Clarke et al. 2007).

Although the testing sessions are fast, it takes several weeks of daily training before the animals learn to locate and grasp the pellets from the steps. Smaller animals can turn around on the platform and retrieve pellets with the opposite forelimb, which conceals a unilateral deficit. Pellet retrieval using the mouth from the upper steps occurs frequently, which may prevent animals from learning the essential grasping motion. However, leaving the upper steps empty during training and testing sessions can easily prevent this unwanted behavior. Lastly, to allow qualitative assessment of reaching actions, each testing box needs to be filmed from two sides, which can place a technical and economic strain on the research group.

2.3 Adhesive removal test

The 'adhesive removal test' is used to measure unilateral tactile extinction and motor deficits. Unlike the previous tests, no food reward is offered in the adhesive removal test; instead, small adhesives are placed on the distal-radial portion of the front paws (Figure 2), which the animals try to remove as quickly as possible. The time taken to contact and to remove the adhesives are recorded (Barth et al. 1990). The adhesive removal test has been used in mouse and rat models of ischemic stroke or TBI to measure sensorimotor impairments (Bouet et al. 2007; Chen et al. 2008; d'Avila et al. 2012; Trueman et al. 2011). After TBI, rats needed significantly longer to notice and contact a small adhesive underneath the front paw contralateral to the lesion compared

to a non-lesioned control group, whereas the time to remove the adhesive from the paw was not affected (Schönfeld et al. 2016).

2.3.1 Testing environment

Testing takes place either in the homecage or in a Plexiglas box; the latter allows filming rats from the bottom, but also requires a habituation period to reduce anxiety to this novel environment.

2.3.2 Testing and analysis of performance

Rats are picked up and an adhesive is attached underneath the left and right paw in a random order. In case of a unilateral lesion, the person attaching the adhesives should be blind to the affected side and apply equal pressure to both paws when positioning the adhesives. The rat will notice (usually by sniffing or shaking the paw) and remove each sticker, one at a time (Barth et al. 1990). A training period before testing is frequently included to reach a stable baseline performance (Bouet et al. 2007); however, based on our experience even untrained, healthy rats manage to remove the adhesives in less than a minute. A trial is finished either after removal of both adhesives or after exceeding a pre-defined time interval and a test session consists of three to five trials, with an interval of five minutes minimum between trials (Bouet et al. 2007; Chen et al. 2008; d'Avila et al. 2012). For each forelimb, the time is recorded until the sticker is contacted and removed and can be reported for the impaired forelimb only or for both forelimbs separately.

2.3.2.1 Standard testing procedure

Adhesives of an equal size are used for both the impaired forelimb and the healthy forelimb and adhesive size stays constant throughout all testing trials.

2.3.2.2 Somatosensory neutralization test

Alternatively, in case of a unilateral lesion the impaired side is identified and adhesive size is increased underneath the impaired limb while it is decreased underneath the healthy limb. Changing the adhesive size is continued until the animal shows an equally fast response for the impaired and healthy forelimb. The ratio of adhesive sizes then indicates the severity of the functional damage (Barth et al. 1990).

2.3.3 Pro and Con

The adhesive removal test can be used to measure two distinct types of behavior: the sensing of a foreign object underneath the paw and its removal. However, in some studies there is no clear distinction made between the two and the latency to remove is the only readout parameter. We discovered that rats with a unilateral cortical lesion showed significant deficits in sensing the adhesive under the impaired paw (Schönfeld et al. 2016) and a strong sensory deficit might be mistaken as motor impairment when both parameters are not separated. Also, adhesive size is not standardized between studies and can vary substantially in diameter; however, no study has been published that describes an influence of adhesive size on reaction time (Chen et al. 2008; d'Avila et al. 2012; Kuypers and Hoane 2010).

2.4 Vermicelli handling test

The 'vermicelli handling test' in rodents is used to assess functional deficits that are comparable to the fine motor skill impairments observed frequently in the

hands of hemiparetic patients. While the animals try to eat food of an unusual shape, goal-directed manipulation with the forelimbs can be observed and the vermicelli handling test has been used to measure deficits in skilled forelimb function in rodent models of TBI, Parkinson's disease and stroke (Allred et al. 2008; Dunkerson et al. 2014; Jacqmain et al. 2014; Plowman et al. 2013). Rats that had received TBI needed significantly more time to eat a given number of pasta pieces compared to healthy controls (Jacqmain et al. 2014).

2.4.1 Testing environment

Animals are provided with pieces of uncooked vermicelli and their handling and eating behavior is observed. Prior to the test sessions, animals should be familiarized with the pasta pieces in their homecage to decrease any fear of novel food (Allred et al. 2008). Also, animals might be food restricted prior to testing to increase their motivation for immediate pasta consumption (d'Avila et al. 2012). The absence (score '0') or presence (score '1') of pasta holding and consumption is scored for each paw in addition to the animal's posture (Allred et al. 2008). Scoring can be done either during the test or offline after test sessions were filmed. Additionally, the number of times an animal drops the pasta piece, the duration of eating and the total number of paw adjustments can be quantified (d'Avila et al. 2012; Dunkerson et al. 2014; Jacqmain et al. 2014).

2.4.2 Pro and Con

Various rating scales allow the experimenter to focus on either coordination of both forelimbs or on paw properties such as digit flexion, thereby creating outcome parameters that are specific to the experimental question but differ

between studies. However, no distinction is made between motor and sensory disabilities (Allred et al. 2008). Care needs to be taken to keep the same brand of pasta and vermicelli pieces of similar length, since rats break down longer pieces prior to handling, which confounds the purpose of the test (Allred et al. 2008).

2.5 Grid test and ladder rung walking task

The grid test and ladder rung walking task assess walking accuracy, coordinated limb use and grasping with fore- and hindlimbs, while the animals traverse an unfamiliar surface with periodically spaced holes. When walking over a grid or a horizontal ladder, animals need to coordinate paw placement and grasp the grid wires or ladder rungs to prevent slipping through the holes. The grid test and ladder rung walking task are frequently used to measure motor impairments before and after mechanical cortical lesions (Adkins et al. 2015; Ajao et al. 2012; Emerick and Kartje 2004; Harris et al. 2010) or stroke (Farr et al. 2006; Ploughman et al. 2007). After TBI, rats showed a strong increase in ‘foot faults’, thus slipping with the paw through the grid, compared to pre-lesion performance (Harris et al. 2010).

2.5.1 Testing environment

2.5.1.1 Grid test

The animals are placed on a grid surface elevated above the ground, with different sizes of grid openings depending on the rodent species tested (Goffus

et al. 2010; Kuypers and Hoane 2010; Onyszchuk et al. 2007). While walking over the grid, the paws occasionally slip off the grid and fall through the wire. In healthy animals those ‘foot faults’ happen equally often with the left and right limb, whereas animals with a unilateral lesion show more foot faults with the contralateral limb (Barth et al. 1990).

2.5.1.2 Ladder rung walking task

The ladder rung walking task consists of two vertical walls that are connected by metal bars that function as ladder rungs. To vary the difficulty of the test, those bars could be spaced at regular or irregular intervals. The animals have to traverse the corridor and while walking over the ladder rungs, they might slip and foot faults occur, that are similar to the deficits observed in the grid test (Metz and Whishaw 2009).

2.5.2 Testing and analysis of performance

Healthy animals readily get habituated to walking over the surface and test sessions consist of counting the number of total steps and foot faults during a period between 2.5 and 5 minutes, either directly during the test or based on video footage (Goffus et al. 2010; Harris et al. 2010; Onyszchuk et al. 2007). Healthy animals cross the grid or ladder fast but animals with a motor deficit might move slower and pause in between, resulting in less total steps during a time-restricted test session. A sufficient number of steps per session can be achieved by recording the walking time instead of ending the session after a fixed time frame (Onyszchuk et al. 2007). Data from the contralateral limb is then shown as a percentage of foot faults relative to all steps.

2.5.2.1 Qualitative analysis

To reveal subtle deficits in paw functioning, a detailed scoring scheme has been developed for the ladder rung walking task to evaluate different severities of foot faults. Based on a frame-by-frame analysis, various types of paw placements can be described on a 7-point rating scale that ranges from 'correct paw placement' to 'total miss'. The individual rating steps are described in detail in Metz and Whishaw (2002).

2.5.3 Pro and Con

No extensive training period is necessary and testing sessions can be conducted quickly. Filming the foot faults with a camera close to the grid or ladder rungs will increase the reliability of the results, since several observers can score the animals. In addition, qualitative analysis of paw placements can help to detect more subtle motor impairments. However, the test cannot be performed with multiple animals in parallel, since continuous attention is necessary to prevent escaping. Some animals may not be willing to move, but their motivation to walk can be stimulated by providing food rewards at the opposite side of the grid.

2.6 Cylinder test

With the 'cylinder test' or 'paw asymmetry test', forelimb use during exploration of a vertical surface can be measured. The cylinder test is based on the inherent curiosity of rodents to explore a new environment and has been used to assess long-term deficits in rats and mice with TBI and ischemic or hemorrhagic stroke (Clarke et al. 2007; Donega et al. 2015; Lam et al. 2013; MacLellan et al. 2013). Rats with a unilateral TBI showed a decreased number of paw contacts with

the cylinder wall using the forelimb that was contralateral to the lesion (Schönfeld et al. 2016).

2.6.1 Testing environment

Animals are placed in a translucent cylinder with a diameter that is smaller than the total length of the animal and a height that prevents escaping (Figure 3). Animals are filmed either from the top, bottom or front, where in the latter case, a mirror should be placed opposite to the camera to film wall contacts made at the back of the cylinder.

2.6.2 Testing and analysis of performance

Explorative activity, which consists of rearing and leaning towards the cylinder wall, is filmed during 5 and 10 minutes. The main readout parameter consists of counting the weight-bearing wall contacts with both forelimbs simultaneously ('both') and the left and right forelimb separately ('left' or 'right'). A wall contact is counted as 'left' or 'right' when both paws had left the cylinder wall before the respective contact or after the rat initiated a new rear (Farr and Trueman 2011; Freret et al. 2009). Either all wall contacts or a predefined number of contacts can be scored and data can be expressed as the percentage of wall contacts with one forelimb relative to the total number of wall contacts (Schallert et al. 2000). In addition, the forelimb used to land after rearing can be noted (Metz et al. 2005).

The cylinder test does not require any form of training; on the contrary, frequent exposure to the cylinders is discouraged since it might decrease the novelty of the environment and therefore explorative behavior. Since the cylinder test

should not be repeated within a short time frame, analysis has to be based on a single test session.

2.6.3 Pro and Con

The cylinder test is used to assess asymmetry between the forelimbs in a single, rather short test session and when the trials are filmed, several animals can be tested in parallel. Repeated sessions in the cylinders do not require an identical testing environment, since novelty stimulates exploration. Still, overexposure to the cylinders through repeated sessions should be avoided. Based on our experience, an interval of two weeks between trials does not decrease vertical exploration behavior.

2.7 Grip strength test

The ‘grip strength test’ measures grasping skills and forelimb muscle strength. The animal must hold on to a bar while it is being pulled away with increasing force in a horizontal direction. With this test, reductions in grip strength resulting from stroke or TBI have been measured in rats and mice (Evans et al. 2015; Linden et al. 2014; Madinier et al. 2014; Maurissen et al. 2003).

2.7.1 Testing equipment

The main equipment needed for the grip strength test is a grasping handle, usually a bar or a ring, connected to a load cell, which measures the force of the handle when being pulled away.

2.7.1.1 Automated grip strength test

The grip strength test can be automated by placing the animal in a chamber with an open slot through which it can access a pull handle on the outside. The pulling force exerted on the handle is then automatically digitized by connected software (Hays et al. 2016).

2.7.2 Behavioral testing and analysis of performance

Animals grasp the handle of the device and are horizontally pulled away until they loose grip. The force needed to break the grip is recorded over 3 to 10 trials giving an indication of forelimb muscle strength. In case of the automated grip strength test, the animal needs to be trained to reach for the pull handle. Training sessions should occur daily for 30 min and after each successful reach, animals are rewarded by sucrose pellets (Hays et al. 2016). Here, the main readout is the force exerted to pull the handle, which in case of a lesion might either be compared to pre-lesion performance or to a fixed threshold (Hays et al. 2016).

2.7.3 Pro and Con

Measuring grip strength is a fast method to determine changes of grasping ability and muscle strength without any specific training of the animal or the experimenter. Animals should be handled with particular care to avoid that they release the handle out of stress or distraction.

A study by Maurissen et al. (2003) pointed out that it is unclear what is measured exactly with the grip strength test, since results can be heavily confounded by several factors, such as the sampling rate of the different devices, the angle from which the tail of the animal is being pulled away and

body weight, which often differs between lesioned animals and healthy controls. Using the automated grip strength test can help to avoid some confounding factors, since the animal does not need to be handled by the experimenter during an ongoing testing session.

2.8 CONCLUDING REMARKS: TESTS TO MEASURE FINE MOTOR SKILLS

Numerous tests are available, which can be used to measure fine motor skill impairments. All tests rely on scoring by an experimenter, which makes them vulnerable to subjective influences. However, quantitative scoring of the single pellet reaching test and the Montoya staircase test may be robust against a biased observer, since counting the number of eaten sucrose pellets is a clearly identifiable readout. When using this type of quantitative assessment (e.g. counting the number of pellets), the Montoya staircase test is superior to the single pellet reaching test in a number of readout parameters that can be measured. A disadvantage of both tests is the duration of training needed for the animals in order to learn how to perform the reaching tasks. Fine motor skill tests that require little to no training of the animals, are the adhesive removal test, the vermicelli handling test, the grid test, the ladder rung walking task, the cylinder test and the grip strength test. The testing sessions are conducted relatively fast and the tests are attractive because they can be performed with equipment available in every lab. However, in each test, additional types of motor behavior are measured: in the adhesive removal test and vermicelli handling test, sensory functions are a prerequisite for forelimb movement, in the grid test and ladder rung walking task, intact gait is a prerequisite and in the

cylinder test, the animals need to keep their balance to be able to rear and stand on the hindlimbs. The grip strength test is used to measure muscle strength in addition to grasping and holding on to a bar. Thus, impairments of sensory functions, balance and gait have to be excluded before interpreting the results of the reviewed tests.

3. TESTS FOR GAIT AND INTER-LIMB COORDINATION

The ability to move in space is vital for all living organisms and gets severely affected when gait is impaired. Gait is the synchronized movement of all limbs in order to move forward and in rodents gait belongs to the most overt behaviors observed in the homecage. Gait is easily observable and therefore, in contrast to fine motor skills, does not require highly specialized tests. Different testing setups are available that range from very basic ones requiring only a flat surface to sophisticated automated systems. In this section, the following tests will be reviewed:

- BBB/BMS: limb placement, joint movements
- CatWalk XT: gait-related limb functions
- Rotarod: walking on a rotating rod
- Beam walking test: crossing an elevated, narrow beam

3.1 Basso, Beattie and Bresnahan (BBB) and Basso Mouse

Scale (BMS) locomotion rating

Developed as a method to evaluate locomotion after spinal cord injury (SCI), the 'Basso, Beattie and Bresnahan' (BBB) scale for rats and the 'Basso mouse scale' (BMS) for mice rely on observer ratings while the animals move around in an open field (Basso et al. 1996; Basso et al. 2006). The BBB consists of a

21-point rating scale, whereas the BMS is reduced to 9 points. Both scales focus on joint movement and coordinated placement of the limbs. Although the BBB has been used after TBI in rats, no functional impairments could be detected. However, when tested using the cylinder test, the rats did display behavioral deficits (Inoue et al. 2013). In the latter study by Inoue et al. (2013) TBI was induced through a precise mechanical impact and the location of the lesion was restricted to the motor cortex. In rodents, the motor cortex is primarily responsible for fine motor skills, such as grasping and handling of small objects. Furthermore, it has been shown that lesions of the motor cortex alone might be insufficient to create impairments of gait and inter-limb coordination (Schönfeld et al. 2016). Given that the BBB measures coordinated paw placements, we suggest that it is not suitable to test motor impairments after focal TBI.

3.1.1 Testing environment

Any flat surface can serve as a basis for the BBB or BMS. Before the test sessions, animals must be habituated to the arena, since open spaces can elicit anxiety in rodents, and the environment must stay constant in case of repeated testing (Basso et al. 1996).

3.1.2 Testing and analysis of performance

After habituation to the open field, animals are left to move around freely for 4 minutes and behavior of the individual limbs is scored directly by at least two blinded observers or it is filmed for offline analysis (Basso et al. 1996). Data is usually shown as the mean score for both fore- or hindlimbs at one timepoint.

3.1.3 Pro and Con

Both, the BMS and BBB scales are widely used to assess locomotor deficits and treatment effects after SCI, which allows for comparisons between different studies. Furthermore, the testing sessions are conducted quickly with minimal equipment. Very subtle changes in limb movements are described in the rating scales; therefore scoring requires experience. To gain experience, observers need to be trained for a considerable amount of time before they can reliably score the animals using either the BMS or the BBB scale. As ratings are done subjectively, at least two blinded observers should score the animals either in parallel or after filming the test session.

3.2 CatWalk XT

The 'CatWalk XT' is an automated analysis system used to assess a large number of parameters associated with gait-related abnormalities and it has been used for rodent models of SCI, muscle denervation, Parkinson's disease and less frequently for stroke or TBI (Chen et al. 2015; Li et al. 2013; Mondello et al. 2015; Mountney et al. 2013; Wang et al. 2011; Westin et al. 2012). In addition to providing data about the spatiotemporal relation between the four paws, functional deficits can be also measured for each individual paw. However, the applicability of the CatWalk XT for animal models with focal, unilateral lesions remains unclear. By using the CatWalk XT, some deficits were detectable in individual limb functioning one month after stroke in rats (Liu et al. 2013; Parkkinen et al. 2013), whereas for rat models of TBI, the measurability of long-term deficits is still under debate (Schönfeld et al. 2016). Mountney et al. (2013) reported a reduction in paw pressure after TBI in rats,

but this was only visible for up to two weeks post lesion, whereas assessment with a 12-point neuroscore scale still showed deficits four weeks after the lesion. In striking contrast, deviations in a broad number of parameters were detectable after TBI in mice for up to one month after the lesion (Cross et al. 2015; Neumann et al. 2009; Wang et al. 2011).

3.2.1 Testing environment

The apparatus consists of a narrow corridor on a glass walkway, which the animal crosses (Figure 4 A-B). The glass walkway is indirectly illuminated in green and lights up upon paw contact. Body contours of the animal are visualized by a red backlight. A camera underneath the walkway films the traversing animal and the software semi-automatically detects footprints and generates a large range of read-out parameters. Examples of such parameters are the mean intensity of the paw prints on the walkway, giving an indication of the pressure that is exerted on the respective paw, the stride length, which is the distance between two placements of the same paw, and walking speed (Schönfeld et al. 2016).

3.2.2 Testing and analysis of performance

Prior to testing, animals must be habituated to the walkway and learn to cross it from end to end, which can be stimulated either by food rewards or a dark goal box at the opposite end (Koopmans et al. 2005; Parkkinen et al. 2013). When opting for food rewards, animals might be food deprived before training and testing sessions.

Each time the animal has crossed the walkway, the software indicates, which runs to keep and discard according to criteria, such as a minimum and maximum run duration and a variation of the walking speed, which should not exceed 30% (Mountney et al. 2013). Furthermore, crossings of the walkway need to be consistent without any interruptions. A minimum of three runs should be obtained per testing session and averaged manually by the experimenter after the gait parameters of each individual run have been extracted.

Based on the filmed runs, the software can assign labels to the four paws (‘left front’, ‘right front’, ‘left hind’, ‘right hind’); however, it is strongly recommended that the experimenter checks the software choices, as they are error-prone. The software might rate runs as suitable for analysis, even if the animal turns in the middle of the walkway and runs back to the direction it originally started. Also labels of ipsilateral front- and hindpaws might be mixed up in the case of spatial overlap.

3.2.3 Pro and Con

For rodent models with overt locomotor impairments, such as models of SCI or Parkinson’s disease, a large amount of data can be obtained using the CatWalk XT. However, the CatWalk XT should not be the only behavioral test in animal models with a unilateral brain lesion, especially when those models lack visible motor impairments (Schönfeld et al. 2016).

Most parameters are influenced by changes in animal weight and running speed, which complicates studies that assess the change of motor deficits over time (Parkkinen et al. 2013). For being an automated test, the involvement of the experimenter is still rather high, since animals have to be trained and

motivated to cross the walkway correctly. When analyzing the runs, choices made by the software should be checked carefully by the experimenter, which increases the amount of time needed for a testing session. As a last note, the CatWalk XT system is quite expensive, which makes it unattainable for many smaller research groups with limited funding resources and therefore reduces number of comparable studies.

3.3 Rotarod test

The 'rotarod' is used to measure the time an animal can stay on a horizontal, rotating rod. However, with this test, it is difficult to distinguish between various gait-related behaviors such as coordination, balance, strength and endurance, which all influence the latency to stay on the rod. In rat and mouse models of TBI, stroke and SCI, animals tend to fall from the rotarod earlier than healthy controls (Brody et al. 2007; Doll et al. 2015; Kim et al. 2015b; Vidal et al. 2013; Vonder Haar et al. 2012).

3.3.1 Testing environment

The rotarod test consists of an elevated rod in different diameters, depending on the rodent species being tested; in general the diameter should be smaller than the length of the tested animal. The rod rotates automatically with various velocities and it can be subdivided into several compartments by plastic divisions, which allows testing of multiple animals at once (Figure 5).

3.3.2 Testing and analysis of performance

3.3.2.1 *Standard testing procedure*

Animals are placed on the slowly rotating rod and speed is increased steadily until the animals fall onto the landing plate (Figure 5). Animals can be habituated to the still rod first and followed by training on the accelerating rod. Testing session consist of two to four trials using conditions similar to the training sessions (Brody et al. 2007; Vonder Haar et al. 2012). Data are afterwards presented as the latency to fall in seconds or as percentages relative to baseline.

3.3.2.2 *Constant speed*

In order to correct for motor abnormalities not directly related to the lesion, latency to fall can be measured at a constant speed (Brody et al. 2007). However, if the constant speed is too slow, differences between injured and healthy animals are not detectable, so it is necessary to choose a rotation speed that is challenging enough for injured animals (Onyszchuk et al. 2007).

3.3.3 Pro and Con

Since the rotarod is used to measure gait in a time-dependent manner, it only requires keeping track of the latency to fall from the rod, which makes the test robust against subjective influences. Furthermore, several animals can be tested at the same time. Unfortunately, the rotarod is a relatively costly device, considering that only a single readout parameter is produced. A decrease in latency to fall can have numerous causes, like coordination deficits or muscle

weakness, which cannot be specified with this testing setup. Based on our own experience, mice may be unwilling to run on the rotarod and may develop the habit of actively jumping off the rod. This creates an artificially short latency to fall, which is unrelated to the severity of any functional impairment. On the contrary, severely impaired mice can develop compensatory strategies to prolong the latency to fall, for example, by tightly gripping on to the rod using just the forelimbs, thereby staying on as the rod rotates. Therefore additional observation of the animals' behavior during the rotarod test is required to exclude those showing undesirable behavior.

3.4 Beam walking test

The 'beam walking test' or 'balance beam test' is used to analyze rodent gait in a testing environment that challenges their ability to maintain balance given that the animals have to cross an elevated beam with a narrow diameter. The beam walking test has been used in rodents to measure impairments after both TBI and stroke (Gao et al. 2015; Qu et al. 2005; Qu et al. 2015). Aged rats that received an experimental stroke showed an increased number of paw slips from the beam with the limb contralateral to the lesion, when compared to control rats of a similar age (Qu et al. 2015).

3.4.1 Testing environment

Depending on the rodent species, animals have to walk over a narrow beam measuring 1-3 cm wide that is elevated between a pole and their homecage or a dark compartment (Chen et al. 2008; Shijo et al. 2015). The material composing the beam can be smooth, made from plastic or wood, or

alternatively a rope can be used, which allows additional support as the animals make use of their claws.

3.4.2 Testing and analysis of performance

3.4.2.1 Standard testing conditions

Training the animals can be useful to achieve a stable baseline measurement (Shijo et al. 2015). The latency to traverse the beam is documented and hindlimb performance can be rated using a 7-point scale (Chen et al. 2008). If animals hesitate to cross the beam, walking can be stimulated by tapping on the tail (Shijo et al. 2015). However, this must be performed in all animals to keep testing sessions comparable.

3.4.2.2 Supporting ledge to prevent compensatory strategies

A ledge can be installed next to the narrow beam to be used as a crutch, on which the impaired limb can be placed (Schallert 2006). With this system in place, healthy animals can then traverse the beam without additional support, whereas animals with impaired fore- or hindlimb function frequently step on the ledge as a support. Using the ledge prevents the development of compensatory strategies, which might conceal a true impairment. As a readout parameter, the number of slips occurring in each limb are documented.

3.4.3 Pro and Con

The beam walking test requires minimal equipment and can be conducted quickly, given that one run should not exceed 60 seconds (Chen et al. 2015). However, scoring occurs during the test period, which make results prone to

subjective influences by the observer. This may be minimized by filming the test sessions and having two blinded observers score the animals.

3.5 CONCLUDING REMARKS: TESTS FOR GAIT AND INTER-LIMB COORDINATION

Several tests are available to quantify gait-related behavior and deviations in limb coordination. As gait is such an overt behavior, tests can be automated, making the reliance on observer-based scores unnecessary. An example of automated gait analysis is the CatWalk XT, where in theory a great amount of data collection and processing can be done by the system without involvement of the experimenter. In practice, the experimenter still needs to invest a considerable amount of time in training the animals and controlling the quality of the acquired data. In addition, automated tests are very expensive, which makes them unattainable for small labs. In contrast to automated systems, the BMS and BBB rely on subjective ratings by an observer, which may be a confounding factor influencing the results. However, using at least two well-trained and blinded observers will reduce these unwanted influences to a minimum. Observer-based rating scales such as the BMS and BBB, based on our experience, have a sensitivity comparable to that of the CatWalk XT. However, to our knowledge a systematic comparison of these tests has not yet been performed.

Unfortunately, the CatWalk XT and the BMS and BBB are time-consuming, because they require training of the animals as well as of the experimenter. The rotarod is an alternative test to measure gait, which requires neither extensive training of the animals (except for a short habituation period), nor experimenter

training. Unfortunately, with the rotarod, gait is evaluated via a single readout parameter that reflects a combination of several motor behaviors. Therefore, the rotarod does not allow discrimination between different types of gait impairment. This is in contrast to the BMS and BBB analyses, which also provide one final outcome parameter but have a more complex scoring scale including subscores which describe additional aspects of gait. Finally, the CatWalk XT produces a wide array of gait-related outcome parameters. However, the biological relevance of each CatWalk XT parameter may not be clear to every investigator.

Similar to the rotarod, the single outcome parameter obtained with the balance beam test does not only represent gait, but also other related behaviors, such as maintaining balance and fine motor skills. However, a potential advantage of the balance beam test over the rotarod is its likely availability in all labs, as the testing environment can be easily built from readily available equipment.

Although it is rather straightforward to measure gait in rodents, it remains challenging to model gait impairments with a similar disease course as observed in the clinical population. Rodents walk on four limbs and recover spontaneously from locomotor impairments through developing compensatory strategies or increased central nervous system plasticity. Therefore, findings that rely on the gait tests reviewed here, should be treated with caution and can rarely be translated directly to human patients.

4. NEUROLOGICAL DEFICIT SCORES

Neurological deficit scores are used to measure different types of motor behavior, providing the experimenter with a single readout score that

summarized all types of impairment. In this section, the following scales for a rodent neurological examination will be reviewed:

- Bederson scale: paw asymmetry and balance, general mobility
- Garcia scale: general mobility, paw asymmetry, grip strength, reaction to sensory stimuli
- Modified neurological severity score: general mobility, reaction to sensory stimuli, paw asymmetry and balance

The Bederson scale and the Garcia scale have been used to detect neurological differences in animals as early as 1-2 days after stroke and the resulting values correlated with the infarct size (Bederson et al. 1986; Garcia et al. 1995). Also, the Bederson scale and the mNSS were used to assess long-term impairment after TBI (Chen et al. 2008; Mountney et al. 2013; Wang et al. 2013).

4.1 Testing and analysis of performance

The equipment needed for the different neurological examinations depends on the type of tests the animal will be subjected to. Concerning the Bederson scale, the necessary equipment is readily available in a lab whereas the Garcia scale requires some simple tools such as a wire mesh, upon which the rat can climb, or a blunt stick such as a cotton swab.

4.1.1 Bederson scale

An observer rates the performance of an animal in distinct tests reflecting different types of neurological functions on a scale from 0-3 during three to five

minutes. When an animal does not show abnormalities in one of the tests, it receives a final score and will not proceed to the next test. Firstly, animals are held at their tails to observe forelimb flexion as a measure of paw asymmetry. Healthy animals extend both forelimbs towards the ground, receiving a score of '0', whereas stroke animals tend to flex the limb contralateral to infarction receiving a score of '1'. Secondly, animals are placed on a plastic-coated paper sheet, gently grabbed by the experimenter and pushed to the side to measure their ability to maintain balance. Healthy animals will resist the lateral push whereas impaired animals show less resistance on the contralateral side (score '2'). Lastly, general mobility, in particular circling behavior, is observed in an open field. Impaired animals show circling towards the paretic side and receive a score of '3'.

4.1.2 Garcia scale

The Garcia scale is subdivided into six tests, which all have to be completed by the animal regardless of the severity of the lesion (Garcia et al. 1995). Impairment is rated with scores that range from 0-3. In test 1, general mobility of the animal in its homecage is observed. In test 2, the animal is held by the tail and symmetry of movement between all four limbs is rated. In test 3, forepaw outstretching as an additional measurement of paw asymmetry is rated. By elevating the tail, the hindlimbs of the animal are lifted upwards and the animal walks using only its forelimbs. In test 4, vertical climbing on a wire mesh is evaluated together with the grip strength when being pulled away from the wire surface. In tests 5 and 6 the animal's reaction to a sensory stimulus (by touching the side of the body and the whiskers) is evaluated.

4.1.3 Modified Neurological Severity Score (mNSS)

The mNSS evaluates neurological functions based on four tests that are comparable to the Garcia scale. In test 1, general mobility is evaluated by raising the animal by its tail and observing locomotion on a flat surface. In test 2, the reaction to sensory stimuli is evaluated by providing visual and tactile sensations and stimulating the hindlimb muscles. In test 3, animals have to maintain their balance while walking over a beam, similar to the beam walking test reviewed earlier and in test 4, asymmetry between the limbs assessed by the presence or absence of reflexes is tested. For a description of the scoring criteria comprising these individual tests, we refer to Chen et al. (2001).

4.2 Pro and Con

Rating neurological deficits on an ordinal scale is a traditional method to define functional deficits in animal models. Unfortunately, fine graduations in disability can be overlooked by using this type of rating scale (Desland et al. 2014). Finally, as with all non-automated tests, rating is based on the observation by an experimenter, which can lead to a score bias.

4.3 CONCLUSION: NEUROLOGICAL DEFICIT SCORES

The different forms of neurological examinations are attractive methods to screen for the presence of a number of behavioral deficits. They can be readily conducted, without costly equipment, but require at least two well-trained and blinded observers to reduce subjective influences during data collection. The different scales used for a neurological examination hamper data comparison between research groups, but within the same research group neurological examinations may be useful to get a quick indication on the neurological status of an animal. Although the scales are complex and measure a broad variety of

rodent behaviors, only one outcome parameter is provided which does not specify the exact type of impairment.

5. GENERAL CONCLUSIONS AND FUTURE PERSPECTIVES

A wide array of tests are available to measure deficits in skilled limb use resulting from stroke, TBI or neurodegenerative diseases (Table 1). In addition to a thorough description of the tests and listing a number of their advantages and disadvantages, we want to provide general recommendations on how to choose the best behavioral test for a specific scientific question (Table 2). Firstly, complex tests are superior to simple tests in the number of investigated aspects that belong to one type of behavior, for example, gait. The BBB assesses gait through rating various individual behaviors, such as hip movements, ankle movements and stepping, whereas the rotarod measures gait only by the time spent on the rotating rod. Simple tests can be readily conducted and are an attractive choice when used as a screening tool for motor impairment; however, even for screening purposes, we prefer to choose complex tests, provided that they are already established in the lab. Secondly, tests which measure multiple readout parameters generate a large amount of data within a single session. During the CatWalk XT, animals have to walk through a corridor and this provides a large amount of gait-related readout parameters, whereas the grid test is used to measure gait and fine motor skills simply by counting the number of foot faults. Thirdly, multiple tests that measure the same underlying behavior, for example fine motor skills of the forelimbs, should be avoided since every test can increase the amount of stress and discomfort experienced by the animals. Fourthly, automated tests are attractive

candidates as they do not rely on subjective, observer-based scores; however, we are aware of their high cost and this may be disadvantageous for small research groups with limited funding. Scoring of neurological deficits needs to be done by at least two blinded and well-trained observers, which makes the testing procedure time consuming as it occupies several members of a research group. Tests such as the CatWalk XT and the rotarod require a minimal contribution from the experimenter while the animals perform the test and therefore, the scores are relatively objective and do not require involvement of more than one experimenter. However, the automated tests reviewed here are only used for gait-related behaviors and cannot measure fine motor skills such as grasping or movement of the wrists. Therefore, we suggest to combine an automated test with a complex fine motor skill test that, if possible, generates multiple readouts, such as the staircase test or the adhesive removal test. The last recommendation may sound self-explanatory but many young researchers fail to follow this simple rule. For investigators who are starting a new research line, it is especially important to choose behavioral tests that are the 'gold standard'. There is no fixed definition of the term 'gold standard' concerning the type of behavioral tests. However, a few rules may help young scientists to identify gold standard tests. Gold standard tests are regularly published for a specific animal model, even if these tests do not fulfill all of the criteria mentioned above. In addition these tests are frequently used by experts in the field (for example experts of a given disease model) and lead to reliable detection of behavioral impairments. Thus, a potential 'gold standard' can be judged by the frequency of publications and the quality of the results.

Usage of the same test across different research groups increases the comparability of results and decreases the number of unnecessary animal experiments. The majority of the above reviewed behavioral tests are suitable to measure motor deficits in rodent models of stroke and TBI. However, comparing the severity of a given impairment between two independent research groups may remain difficult, since timepoints of assessment, scoring schedule and data presentation can vary greatly (Table 1). Additional complications might be created by ordinal rating scales that are used in complex tests such as the BMS, the BBB or the neurological deficit scores. When using an ordinal rating scale, a two-fold increase in scores does not mean that one animal performs twice as well as another animal, as the individual parts of the respective tests measure different functions, but the final performance is summarized with one score.

Overall, there is a strong need for a consensus on how to perform animal experiments and present data in a uniform way (Lemmon et al. 2014). When measuring unilateral functional impairments after a cortical lesion, data are not always presented as functional scores of the impaired limb, but sometimes differences or ratios between the impaired and the healthy limb are shown instead. To make data presentation more comparable, a combination of results can be shown (Figure 6): One graph consisting of data of the healthy forelimb, one graph consisting of data of the impaired forelimb and additionally a 'difference score' calculated by subtracting the score of the healthy limb from the score of the impaired limb as a direct measure of limb asymmetry.

5.1 The effects of behavioral tests on animals: Stress or a stimulating variety?

Planning and performing animal experiments must be done according to the 3Rs: replacement, reduction and refinement (Manciocco et al. 2009). To answer scientific questions related to locomotion or specific motor skills, animal models are used since alternative methods remain unavailable. These animal models should represent key characteristics of the clinical population as close as possible, for example in the case of TBI by creating brain lesions through mechanical trauma and by including male animals, since the occurrence of TBI is highest amongst males. Nevertheless, the experimental paradigm can be refined and the number of animals in an experiment can be reduced by carefully choosing the appropriate behavioral tests. Table 2 summarizes criteria to choose the best behavioral test for a new research line.

It is clear that execution of behavioral tests leads to an interruption of an animal's daily life in a cage; the question remains whether this interruption is stressful or stimulating. Handling by the experimenter and the introduction to a foreign environment can elicit stress. Furthermore, the specific design of some testing setups can trigger anxiety or exacerbate pre-existing pain. Examples include placing animals in a large open field (Prut and Belzung 2003) to allow manual scoring of motor behavior with the BBB or BMS or introducing the animals to small testing boxes (e.g. Montoya staircase test). Schallert et al. (2003) suggest that animals get habituated to the experimenter and the testing paradigms after gentle and repeated exposure to the new sensations. Habituation may decrease the experienced stress levels drastically; however, this needs to be supported by systematic analyses of physiological stress parameters, such as blood corticosterone levels.

Some testing paradigms require additional procedures such as food restriction, which creates a stressful component that does not decrease after repeated exposure and may have a negative impact on the results. Therefore, it remains debatable whether testing paradigms that require food restriction are acceptable or should be replaced by paradigms that measure comparable behavioral functions, without the need to interfere with the feeding behavior of the animal.

We provide an extensive overview of several rodent motor skill tests that fall into three distinct categories, based on the type of motor function assessed. None of the categories contains tests that are systematically more prone to cause discomfort to the animals than tests from a different category; therefore no general ethical recommendation for a specific test (category) can be provided. Researchers need to decide based on their research question, which categories of motor skill tests are suited best and then, if possible, choose specific tests that do not involve the potential stressors mentioned above. On the other hand, it would be ethically questionable to perform animal experiments involving a behavioral test that guarantees the absence of any discomfort, but does not reliably measure a given type of motor functions. Therefore, costs and benefits of animal experiments have to be weighted carefully. In general the level of discomfort caused by behavioral tests needs to be kept as low as possible in order to comply with the 3Rs and to increase the quality of the behavioral data.

5.2 Sensory deficits and pain might worsen motor impairment

Due to a close spatial proximity between the motor cortex and the somatosensory cortex in rodents (Benali et al. 2008) and the high amount of connectivity between both regions (Farkas et al. 1999), lesions to the motor cortex may spread towards surrounding cortical areas and cause a significant dysfunction in sensory regions. Furthermore, chronic pain can occur after lesions to the central nervous system and poses a well-known complication after SCI or stroke in humans (Balazy 1992; Mulla et al. 2015). Both sensory deficits and pain can negatively influence motor behavior in rodents; for example, an animal may not grasp a food pellet because it does not sense it or if the movement creates pain, although the locomotor functions are intact. To detect additional sensory impairments the 'corner test' or the 'vibrissae-elicited forelimb placing test' may be useful (Bouet et al. 2007; Kuypers and Hoane 2010). Tests to detect pain-related behavior are reviewed elsewhere in detail (Barrot 2012). Thus, if there is any indication that sensory deficits and/or pain may influence the analysis of the motor functions, these tests should be performed.

5.3 Concluding remarks

Measuring functional impairment in rodent models of stroke or TBI is crucial to understand the course of the disease and the efficacy of novel treatments. The chosen tests should be sensitive enough to detect subtle behavioral impairments over an extended period of time and document various characteristics of the deficit, while causing minimal discomfort. The current review summarizes a number of tests, which fulfill those criteria (Table 1). This

can be used as a guide for researchers to select the optimal tests for measuring skilled limb use in animal models with unilateral motor deficits (Table 2).

Acknowledgements

This work was supported in part by a grant from Hasselt University (BOF11603BOFMO) to SH and a grant from the Hersenstichting Nederland to YT.

Author disclosure statement

No competing financial interests exist.

REFERENCES

- Abrous DN, Dunnett SB (1994) Paw reaching in rats: the staircase test
Neuroscience Protocols 3:1-11
- Adkins DL et al. (2015) Combining Multiple Types of Motor Rehabilitation
Enhances Skilled Forelimb Use Following Experimental Traumatic Brain
Injury in Rats Neurorehabilitation and neural repair
doi:10.1177/1545968315576577
- Ajao DO et al. (2012) Traumatic brain injury in young rats leads to progressive
behavioral deficits coincident with altered tissue properties in adulthood
Journal of neurotrauma 29:2060-2074 doi:10.1089/neu.2011.1883
- Allred RP et al. (2008) The vermicelli handling test: a simple quantitative
measure of dexterous forepaw function in rats Journal of neuroscience
methods 170:229-244 doi:10.1016/j.jneumeth.2008.01.015
- Baird AL, Meldrum A, Dunnett SB (2001) The staircase test of skilled reaching
in mice Brain research bulletin 54:243-250
- Balazy TE (1992) Clinical management of chronic pain in spinal cord injury The
Clinical journal of pain 8:102-110
- Barrot M (2012) Tests and models of nociception and pain in rodents
Neuroscience 211:39-50 doi:10.1016/j.neuroscience.2011.12.041
- Barth TM, Jones TA, Schallert T (1990) Functional subdivisions of the rat
somatic sensorimotor cortex Behavioural brain research 39:73-95
- Basso DM, Beattie MS, Bresnahan JC (1996) Graded histological and
locomotor outcomes after spinal cord contusion using the NYU weight-
drop device versus transection Experimental neurology 139:244-256
doi:10.1006/exnr.1996.0098

- Basso DM, Fisher LC, Anderson AJ, Jakeman LB, McTigue DM, Popovich PG (2006) Basso Mouse Scale for locomotion detects differences in recovery after spinal cord injury in five common mouse strains *Journal of neurotrauma* 23:635-659 doi:10.1089/neu.2006.23.635
- Beaudin SA, Nisam S, Smith DR (2013) Early life versus lifelong oral manganese exposure differently impairs skilled forelimb performance in adult rats *Neurotoxicology and teratology* 38:36-45 doi:10.1016/j.ntt.2013.04.004
- Bederson JB, Pitts LH, Tsuji M, Nishimura MC, Davis RL, Bartkowski H (1986) Rat middle cerebral artery occlusion: evaluation of the model and development of a neurologic examination *Stroke; a journal of cerebral circulation* 17:472-476
- Benali A, Weiler E, Benali Y, Dinse HR, Eysel UT (2008) Excitation and inhibition jointly regulate cortical reorganization in adult rats *The Journal of neuroscience : the official journal of the Society for Neuroscience* 28:12284-12293 doi:10.1523/JNEUROSCI.1952-08.2008
- Bouet V, Freret T, Toutain J, Divoux D, Boulouard M, Schumann-Bard P (2007) Sensorimotor and cognitive deficits after transient middle cerebral artery occlusion in the mouse *Experimental neurology* 203:555-567 doi:10.1016/j.expneurol.2006.09.006
- Brody DL et al. (2007) Electromagnetic controlled cortical impact device for precise, graded experimental traumatic brain injury *Journal of neurotrauma* 24:657-673 doi:10.1089/neu.2006.0011

- Carmichael ST, Wei L, Rovainen CM, Woolsey TA (2001) New patterns of intracortical projections after focal cortical stroke *Neurobiology of disease* 8:910-922 doi:10.1006/nbdi.2001.0425
- Cenci MA, Whishaw IQ, Schallert T (2002) Animal models of neurological deficits: how relevant is the rat? *Nature reviews Neuroscience* 3:574-579 doi:10.1038/nrn877
- Chen CJ et al. (2015) Improved neurological outcome by intramuscular injection of human amniotic fluid derived stem cells in a muscle denervation model *PLoS One* 10:e0124624 doi:10.1371/journal.pone.0124624
- Chen J, Li Y, Wang L, Zhang Z, Lu D, Lu M, Chopp M (2001) Therapeutic benefit of intravenous administration of bone marrow stromal cells after cerebral ischemia in rats *Stroke; a journal of cerebral circulation* 32:1005-1011
- Chen SF, Hsu CW, Huang WH, Wang JY (2008) Post-injury baicalein improves histological and functional outcomes and reduces inflammatory cytokines after experimental traumatic brain injury *British journal of pharmacology* 155:1279-1296 doi:10.1038/bjp.2008.345
- Clarke J, Ploughman M, Corbett D (2007) A qualitative and quantitative analysis of skilled forelimb reaching impairment following intracerebral hemorrhage in rats *Brain research* 1145:204-212 doi:10.1016/j.brainres.2007.01.135
- Cross DJ et al. (2015) Paclitaxel improves outcome from traumatic brain injury *Brain research* 1618:299-308 doi:10.1016/j.brainres.2015.06.006

- d'Avila JC et al. (2012) Microglial activation induced by brain trauma is suppressed by post-injury treatment with a PARP inhibitor *Journal of neuroinflammation* 9:31 doi:10.1186/1742-2094-9-31
- Desland FA, Afzal A, Warraich Z, Mocco J (2014) Manual versus Automated Rodent Behavioral Assessment: Comparing Efficacy and Ease of Bederson and Garcia Neurological Deficit Scores to an Open Field Video-Tracking System *Journal of central nervous system disease* 6:7-14 doi:10.4137/JCNSD.S13194
- Doll DN, Engler-Chiurazzi EB, Lewis SE, Hu H, Kerr AE, Ren X, Simpkins JW (2015) Lipopolysaccharide exacerbates infarct size and results in worsened post-stroke behavioral outcomes *Behavioral and brain functions* : BBF 11:32 doi:10.1186/s12993-015-0077-5
- Donega V et al. (2015) Assessment of long-term safety and efficacy of intranasal mesenchymal stem cell treatment for neonatal brain injury in the mouse *Pediatric research* 78:520-526 doi:10.1038/pr.2015.145
- Dunkerson J et al. (2014) Combining enriched environment and induced pluripotent stem cell therapy results in improved cognitive and motor function following traumatic brain injury *Restorative neurology and neuroscience* 32:675-687 doi:10.3233/RNN-140408
- Emerick AJ, Kartje GL (2004) Behavioral recovery and anatomical plasticity in adult rats after cortical lesion and treatment with monoclonal antibody IN-1 *Behavioural brain research* 152:315-325 doi:10.1016/j.bbr.2003.10.010
- Evans TM, Jaramillo CA, Sataranatarajan K, Watts L, Sabia M, Qi W, Van Remmen H (2015) The effect of mild traumatic brain injury on peripheral

- nervous system pathology in wild-type mice and the G93A mutant mouse model of motor neuron disease *Neuroscience* 298:410-423 doi:10.1016/j.neuroscience.2015.04.041
- Farkas T, Kis Z, Toldi J, Wolff JR (1999) Activation of the primary motor cortex by somatosensory stimulation in adult rats is mediated mainly by associational connections from the somatosensory cortex *Neuroscience* 90:353-361
- Farr TD, Liu L, Colwell KL, Whishaw IQ, Metz GA (2006) Bilateral alteration in stepping pattern after unilateral motor cortex injury: a new test strategy for analysis of skilled limb movements in neurological mouse models *Journal of neuroscience methods* 153:104-113 doi:10.1016/j.jneumeth.2005.10.011
- Farr TD, Trueman RC (2011) Functional assessment of subcortical ischemia. In: *Animal models of movement disorders, vol II. Neuromethods, vol 62.* Springer Science+Business Media, Berlin, Germany, pp 91-114. doi:DOI 10.1007
- Freret T et al. (2009) Behavioral deficits after distal focal cerebral ischemia in mice: Usefulness of adhesive removal test *Behavioral neuroscience* 123:224-230 doi:10.1037/a0014157
- Gao Y et al. (2015) Mice lacking glutamate carboxypeptidase II develop normally, but are less susceptible to traumatic brain injury *Journal of neurochemistry* doi:10.1111/jnc.13123
- Garcia JH, Wagner S, Liu KF, Hu XJ (1995) Neurological deficit and extent of neuronal necrosis attributable to middle cerebral artery occlusion in rats.

Statistical validation Stroke; a journal of cerebral circulation 26:627-634;
discussion 635

Gilmour G, Iversen SD, O'Neill MF, Bannerman DM (2004) The effects of
intracortical endothelin-1 injections on skilled forelimb use: implications
for modelling recovery of function after stroke Behavioural brain
research 150:171-183 doi:10.1016/j.bbr.2003.07.006

Giri S, Bader A (2015) A low-cost, high-quality new drug discovery process
using patient-derived induced pluripotent stem cells Drug discovery
today 20:37-49 doi:10.1016/j.drudis.2014.10.011

Goffus AM, Anderson GD, Hoane M (2010) Sustained delivery of nicotinamide
limits cortical injury and improves functional recovery following traumatic
brain injury Oxidative medicine and cellular longevity 3:145-152
doi:10.4161/oxim.3.2.11315

Harris NG, Mironova YA, Hovda DA, Sutton RL (2010) Chondroitinase ABC
enhances pericontusion axonal sprouting but does not confer robust
improvements in behavioral recovery Journal of neurotrauma 27:1971-
1982 doi:10.1089/neu.2010.1470

Hays SA, Ruiz A, Bethea T, Khodaparast N, Carmel JB, Rennaker RL, 2nd,
Kilgard MP (2016) Vagus nerve stimulation during rehabilitative training
enhances recovery of forelimb function after ischemic stroke in aged rats
Neurobiology of aging 43:111-118
doi:10.1016/j.neurobiolaging.2016.03.030

Inoue T et al. (2013) Combined SCI and TBI: recovery of forelimb function after
unilateral cervical spinal cord injury (SCI) is retarded by contralateral
traumatic brain injury (TBI), and ipsilateral TBI balances the effects of

- SCI on paw placement *Experimental neurology* 248:136-147
doi:10.1016/j.expneurol.2013.06.006
- Jacqmain J, Nudi ET, Fluharty S, Smith JS (2014) Pre and post-injury environmental enrichment effects functional recovery following medial frontal cortical contusion injury in rats *Behavioural brain research* 275:201-211 doi:10.1016/j.bbr.2014.08.056
- Kim J, Lee H, Selimovic S, Gauvin R, Bae H (2015a) Organ-on-a-chip: development and clinical prospects toward toxicity assessment with an emphasis on bone marrow *Drug safety* 38:409-418 doi:10.1007/s40264-015-0284-x
- Kim K et al. (2015b) The Effect of Human Umbilical Cord Blood-Derived Mesenchymal Stem Cells in a Collagenase-Induced Intracerebral Hemorrhage Rat Model *Experimental neurobiology* 24:146-155 doi:10.5607/en.2015.24.2.146
- Kloth V, Klein A, Loettrich D, Nikkhah G (2006) Colour-coded pellets increase the sensitivity of the staircase test to differentiate skilled forelimb performances of control and 6-hydroxydopamine lesioned rats *Brain research bulletin* 70:68-80 doi:10.1016/j.brainresbull.2006.04.006
- Koopmans GC, Deumens R, Honig WM, Hamers FP, Steinbusch HW, Joosten EA (2005) The assessment of locomotor function in spinal cord injured rats: the importance of objective analysis of coordination *Journal of neurotrauma* 22:214-225 doi:10.1089/neu.2005.22.214
- Kuypers NJ, Hoane MR (2010) Pyridoxine administration improves behavioral and anatomical outcome after unilateral contusion injury in the rat *Journal of neurotrauma* 27:1275-1282 doi:10.1089/neu.2010.1327

- Lai S, Panarese A, Spalletti C, Alia C, Ghionzoli A, Caleo M, Micera S (2015) Quantitative kinematic characterization of reaching impairments in mice after a stroke *Neurorehabilitation and neural repair* 29:382-392 doi:10.1177/1545968314545174
- Lam TI et al. (2013) Beneficial effects of minocycline and botulinum toxin-induced constraint physical therapy following experimental traumatic brain injury *Neurorehabilitation and neural repair* 27:889-899 doi:10.1177/1545968313491003
- Lemmon VP et al. (2014) Minimum information about a spinal cord injury experiment: a proposed reporting standard for spinal cord injury experiments *Journal of neurotrauma* 31:1354-1361 doi:10.1089/neu.2014.3400
- Li S et al. (2013) Assessing gait impairment after permanent middle cerebral artery occlusion in rats using an automated computer-aided control system *Behavioural brain research* 250:174-191 doi:10.1016/j.bbr.2013.04.044
- Linden J, Fassotte L, Tirelli E, Plumier JC, Ferrara A (2014) Assessment of behavioral flexibility after middle cerebral artery occlusion in mice *Behavioural brain research* 258:127-137 doi:10.1016/j.bbr.2013.10.028
- Liu Y et al. (2013) Quantitative gait analysis of long-term locomotion deficits in classical unilateral striatal intracerebral hemorrhage rat model *Behavioural brain research* 257:166-177 doi:10.1016/j.bbr.2013.10.007
- Liu Z et al. (2014) Beneficial effects of gfap/vimentin reactive astrocytes for axonal remodeling and motor behavioral recovery in mice after stroke *Glia* 62:2022-2033 doi:10.1002/glia.22723

- MacLellan CL, Langdon KD, Botsford A, Butt S, Corbett D (2013) A model of persistent learned nonuse following focal ischemia in rats Neurorehabilitation and neural repair 27:900-907 doi:10.1177/1545968313496323
- Madinier A, Wieloch T, Olsson R, Ruscher K (2014) Impact of estrogen receptor beta activation on functional recovery after experimental stroke Behavioural brain research 261:282-288 doi:10.1016/j.bbr.2013.12.046
- Manciocco A, Chiarotti F, Vitale A, Calamandrei G, Laviola G, Alleva E (2009) The application of Russell and Burch 3R principle in rodent models of neurodegenerative disease: the case of Parkinson's disease Neuroscience and biobehavioral reviews 33:18-32 doi:10.1016/j.neubiorev.2008.08.002
- Maurissen JP, Marable BR, Andrus AK, Stebbins KE (2003) Factors affecting grip strength testing Neurotoxicology and teratology 25:543-553
- Metz GA, Antonow-Schlorke I, Witte OW (2005) Motor improvements after focal cortical ischemia in adult rats are mediated by compensatory mechanisms Behavioural brain research 162:71-82 doi:10.1016/j.bbr.2005.03.002
- Metz GA, Whishaw IQ (2002) Drug-induced rotation intensity in unilateral dopamine-depleted rats is not correlated with end point or qualitative measures of forelimb or hindlimb motor performance Neuroscience 111:325-336
- Metz GA, Whishaw IQ (2009) The ladder rung walking task: a scoring system and its practical application Journal of visualized experiments : JoVE doi:10.3791/1204

- Mondello SE, Sunshine MD, Fishedick AE, Moritz CT, Horner PJ (2015) A cervical hemi-contusion spinal cord injury model for the investigation of novel therapeutics targeting proximal and distal forelimb functional recovery *Journal of neurotrauma* doi:10.1089/neu.2014.3792
- Montoya CP, Campbell-Hope LJ, Pemberton KD, Dunnett SB (1991) The "staircase test": a measure of independent forelimb reaching and grasping abilities in rats *Journal of neuroscience methods* 36:219-228
- Mountney A, Leung LY, Pedersen R, Shear D, Tortella F (2013) Longitudinal assessment of gait abnormalities following penetrating ballistic-like brain injury in rats *Journal of neuroscience methods* 212:1-16 doi:10.1016/j.jneumeth.2012.08.025
- Mulla SM et al. (2015) Management of Central Poststroke Pain: Systematic Review of Randomized Controlled Trials *Stroke; a journal of cerebral circulation* 46:2853-2860 doi:10.1161/STROKEAHA.115.010259
- Nakagawa H, Ueno M, Itokazu T, Yamashita T (2013) Bilateral movement training promotes axonal remodeling of the corticospinal tract and recovery of motor function following traumatic brain injury in mice *Cell death & disease* 4:e534 doi:10.1038/cddis.2013.62
- Neumann M, Wang Y, Kim S, Hong SM, Jeng L, Bilgen M, Liu J (2009) Assessing gait impairment following experimental traumatic brain injury in mice *Journal of neuroscience methods* 176:34-44 doi:10.1016/j.jneumeth.2008.08.026
- Onyszchuk G, Al-Hafez B, He YY, Bilgen M, Berman NE, Brooks WM (2007) A mouse model of sensorimotor controlled cortical impact: characterization using longitudinal magnetic resonance imaging, behavioral

- assessments and histology *Journal of neuroscience methods* 160:187-196 doi:10.1016/j.jneumeth.2006.09.007
- Pagnussat Ade S, Michaelsen SM, Achaval M, Netto CA (2009) Skilled forelimb reaching in Wistar rats: evaluation by means of Montoya staircase test *Journal of neuroscience methods* 177:115-121 doi:10.1016/j.jneumeth.2008.10.001
- Parkkinen S, Ortega FJ, Kuptsova K, Huttunen J, Tarkka I, Jolkkonen J (2013) Gait impairment in a rat model of focal cerebral ischemia *Stroke research and treatment* 2013:410972 doi:10.1155/2013/410972
- Ploughman M, Attwood Z, White N, Dore JJ, Corbett D (2007) Endurance exercise facilitates relearning of forelimb motor skill after focal ischemia *The European journal of neuroscience* 25:3453-3460 doi:10.1111/j.1460-9568.2007.05591.x
- Plowman EK et al. (2013) Differential sensitivity of cranial and limb motor function to nigrostriatal dopamine depletion *Behavioural brain research* 237:157-163 doi:10.1016/j.bbr.2012.09.031
- Prescott MJ (2010) Ethics of primate use *Advances in Science & Research* 5:11-22
- Prut L, Belzung C (2003) The open field as a paradigm to measure the effects of drugs on anxiety-like behaviors: a review *European journal of pharmacology* 463:3-33
- Qu C, Lu D, Goussev A, Schallert T, Mahmood A, Chopp M (2005) Effect of atorvastatin on spatial memory, neuronal survival, and vascular density in female rats after traumatic brain injury *Journal of neurosurgery* 103:695-701 doi:10.3171/jns.2005.103.4.0695

- Qu HL et al. (2015) Forced limb-use enhanced neurogenesis and behavioral recovery after stroke in the aged rats *Neuroscience* 286:316-324 doi:10.1016/j.neuroscience.2014.11.040
- Russel WMS, Burch RL (1959) *The Principles of Humane Experimental Technique*. Methuen, London
- Santos MV, Pagnussat AS, Mestriner RG, Netto CA (2013) Motor Skill Training Promotes Sensorimotor Recovery and Increases Microtubule-Associated Protein-2 (MAP-2) Immunoreactivity in the Motor Cortex after Intracerebral Hemorrhage in the Rat *ISRN neurology* 2013:159184 doi:10.1155/2013/159184
- Schallert T (2006) Behavioral tests for preclinical intervention assessment *NeuroRx : the journal of the American Society for Experimental NeuroTherapeutics* 3:497-504 doi:10.1016/j.nurx.2006.08.001
- Schallert T, Fleming SM, Leasure JL, Tillerson JL, Bland ST (2000) CNS plasticity and assessment of forelimb sensorimotor outcome in unilateral rat models of stroke, cortical ablation, parkinsonism and spinal cord injury *Neuropharmacology* 39:777-787
- Schallert T, Woodlee MT, Fleming SM (2003) Experimental focal ischemic injury: behavior-brain interactions and issues of animal handling and housing *ILAR journal / National Research Council, Institute of Laboratory Animal Resources* 44:130-143
- Schönfeld LM et al. (2016) Long-Term Motor Deficits after Controlled Cortical Impact in Rats Can Be Detected by Fine Motor Skill Tests but Not by Automated Gait Analysis *Journal of neurotrauma* doi:10.1089/neu.2016.4440

- Shijo K, Ghavim S, Harris NG, Hovda DA, Sutton RL (2015) Glucose administration after traumatic brain injury exerts some benefits and no adverse effects on behavioral and histological outcomes *Brain research* 1614:94-104 doi:10.1016/j.brainres.2015.04.022
- Starkey ML, Bleul C, Zorner B, Lindau NT, Mueggler T, Rudin M, Schwab ME (2012) Back seat driving: hindlimb corticospinal neurons assume forelimb control following ischaemic stroke *Brain : a journal of neurology* 135:3265-3281 doi:10.1093/brain/aws270
- Tannenbaum J, Bennett BT (2015) Russell and Burch's 3Rs then and now: the need for clarity in definition and purpose *Journal of the American Association for Laboratory Animal Science : JAALAS* 54:120-132
- Trueman RC, Harrison DJ, Dwyer DM, Dunnett SB, Hoehn M, Farr TD (2011) A Critical Re-Examination of the Intraluminal Filament MCAO Model: Impact of External Carotid Artery Transection *Translational stroke research* 2:651-661 doi:10.1007/s12975-011-0102-4
- Vaagenes IC, Tsai SY, Ton ST, Husak VA, McGuire SO, O'Brien TE, Kartje GL (2015) Binge ethanol prior to traumatic brain injury worsens sensorimotor functional recovery in rats *PLoS One* 10:e0120356 doi:10.1371/journal.pone.0120356
- Vidal PM, Lemmens E, Geboes L, Vangansewinkel T, Nelissen S, Hendrix S (2013) Late blocking of peripheral TNF-alpha is ineffective after spinal cord injury in mice *Immunobiology* 218:281-284 doi:10.1016/j.imbio.2012.05.007
- Vonder Haar C, Emery MA, Hoane MR (2012) Chronic folic acid administration confers no treatment effects in either a high or low dose following

- unilateral controlled cortical impact injury in the rat Restorative neurology and neuroscience 30:291-302 doi:10.3233/RNN-2012-110196
- Wang S et al. (2013) Comparison of neurite density measured by MRI and histology after TBI PLoS One 8:e63511 doi:10.1371/journal.pone.0063511
- Wang Y, Neumann M, Hansen K, Hong SM, Kim S, Noble-Haeusslein LJ, Liu J (2011) Fluoxetine increases hippocampal neurogenesis and induces epigenetic factors but does not improve functional recovery after traumatic brain injury Journal of neurotrauma 28:259-268 doi:10.1089/neu.2010.1648
- Westin JE, Janssen ML, Sager TN, Temel Y (2012) Automated gait analysis in bilateral parkinsonian rats and the role of L-DOPA therapy Behavioural brain research 226:519-528 doi:10.1016/j.bbr.2011.10.006
- Whishaw IQ (2000) Loss of the innate cortical engram for action patterns used in skilled reaching and the development of behavioral compensation following motor cortex lesions in the rat Neuropharmacology 39:788-805
- Windle V, Szymanska A, Granter-Button S, White C, Buist R, Peeling J, Corbett D (2006) An analysis of four different methods of producing focal cerebral ischemia with endothelin-1 in the rat Experimental neurology 201:324-334 doi:10.1016/j.expneurol.2006.04.012

FIGURE CAPTIONS

Figure 1: A rat is shown inside of the Montoya staircase testing box, reaching towards the staircase and grasping a red sucrose pellet (arrow). Sucrose pellets lying on each step were color-coded in blue, red, green, purple and yellow (arrowheads) to allow the collection of additional information on grasping success. Note that the upper two steps were left empty; from these steps, the rats could retrieve pellets by mouth, which prevents grasping behavior.

Figure 2: The ability to feel and remove adhesives from the paws is tested with the adhesive removal test. Here, a blue sticker dot was placed underneath the left front paw (black arrow) and a yellow sticker dot was placed underneath the right front paw (white arrow). The animal's reaction towards these stickers is filmed with a camera located underneath a transparent box.

Figure 3: The cylinder test measures exploration of vertical surfaces. The rat shown in the picture rears and leans towards the cylinder wall using the forelimbs (red circles). Counting the number of wall contact allows detection of a potential asymmetry between the forelimbs.

Figure 4: The CatWalk XT is an automated gait analysis system consisting of a walkway (arrow 1) situated above a camera (arrow 2; A). The rat has crossed the glass plate at the bottom of the walkway from one end to the other and waits for a reward provided by the experimenter (B).

Figure 5: The rotarod measures the latency of an animal to stay on a rotating rod. Several mice can walk on the rod in parallel and the rotating velocity is increased until the animals slip (black arrow) and fall down from the rod on the landing plate (white arrow).

Upon contact with the landing plate, the latency to stay on the rod is stopped automatically for each animal.

Figure 6: To increase comparability of results between studies, the following way of standardized data presentation is proposed: One graph showing the functional score of the healthy limb ('A'), ipsilateral to the lesion, and one graph showing the functional score of the impaired limb ('B'), contralateral to the lesion. In addition, the so-called 'difference score' is calculated by subtracting the score of the healthy limb from the score of the impaired limb ('B-A').

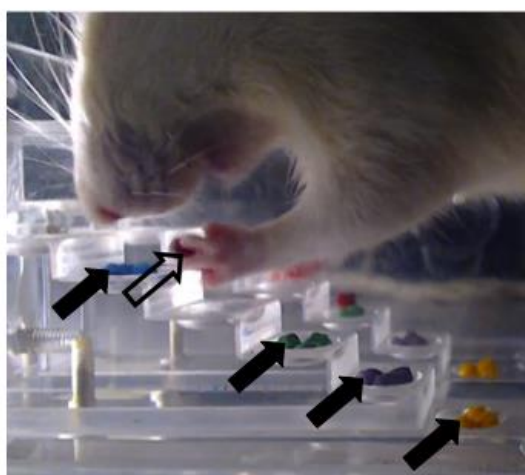


Fig. 1

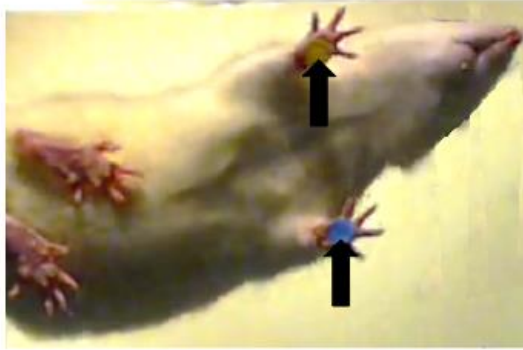


Fig. 2



Fig. 3

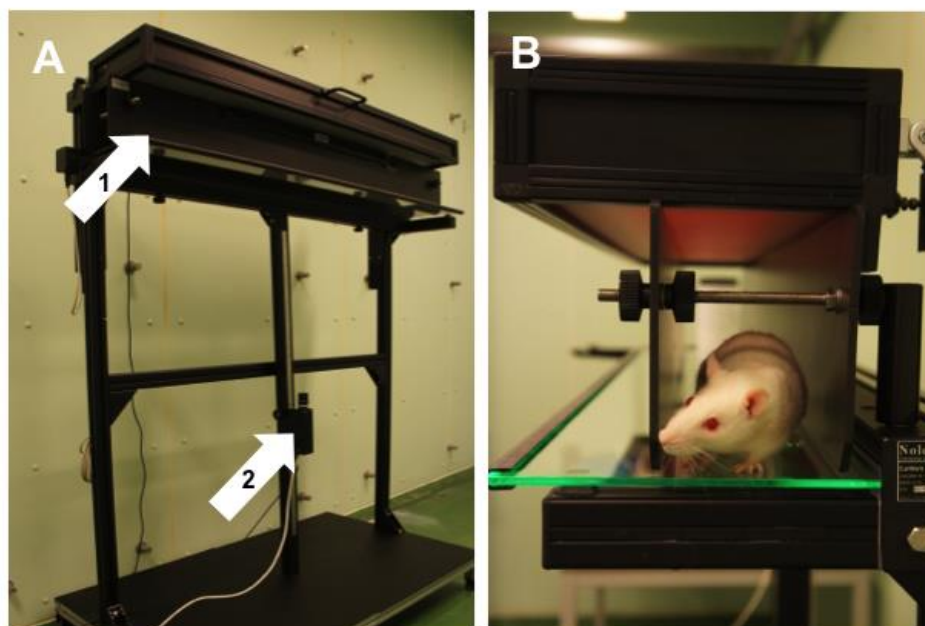


Fig. 4

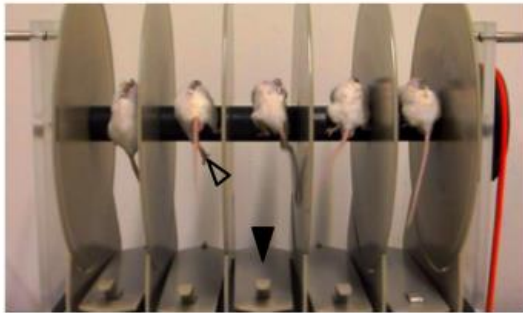


Fig. 5

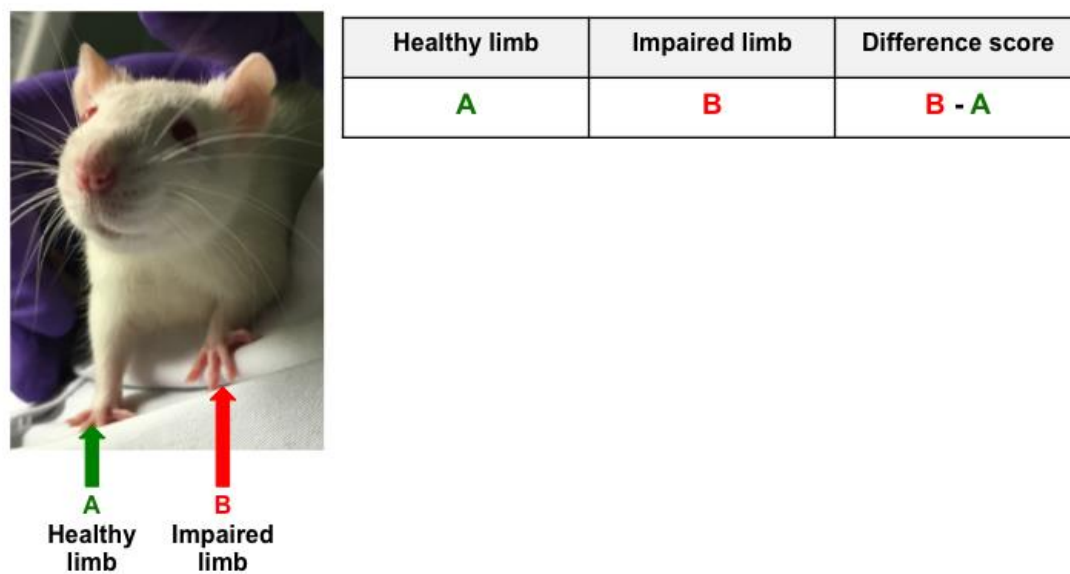


Fig. 6

TABLE CAPTIONS

Table 1: An overview of the reviewed tests including the respective animal models, the different testing paradigms, the readout parameters and the way of data presentation. The mentioned animal models refer to both rats and mice if not stated otherwise, the number of weeks stated in brackets indicates the duration of assessment done in the cited studies. TBI: traumatic brain injury; SCI: spinal cord injury; n/a: not applicable; w: weeks.

TESTS TO MEASURE FINE MOTOR SKILLS					
Test	Animal models	Behavioral assessment	Readouts	Data presentation	References
Single pellet reaching test	TBI (6 w.) Stroke (6 w.)	Pellet presentation to impaired side Qualitative: Limb movement, body position	Grasped and eaten Grasped but dropped	Success rate: Retrieval/trials of both limbs or impaired limb only Qualitative: Various rating scales	Adkins <i>et al.</i> , 2015 Starkey <i>et al.</i> , 2012 Whishaw, 2000
Staircase test	TBI (6 w.) Stroke (4 w.)	Both sides baited Side test	Total eaten pellets per side	Number of pellets at side of contralesional limb or both sides	Clarke <i>et al.</i> , 2007 Kloth <i>et al.</i> , 2006

	Parkinson's disease (6 w.) Peripheral nerve lesion (2 w.)	Speed test Color coding Qualitative rating	Pellets remaining, misplaced, lost, taken and eaten Pellets retrieved, dropped, left Qualitative scores of reaching action	Qualitative: Scores 0-1 for 13 typical reaching actions	Montoya <i>et al.</i> , 1991 Pagnussat <i>et al.</i> , 2009 Schönfeld <i>et al.</i> , 2016
Adhesive removal test	TBI (6 w.) Stroke (4 w.)	Fixed adhesive size Somatosensory neutralization test	Time to contact Time to remove	Time in seconds for both limbs or impaired limb only	Bouet <i>et al.</i> , 2007 Schönfeld <i>et al.</i> , 2016
Vermicelli handling	TBI (8 w.) Stroke (6 w.) Parkinson's disease (6 w.)	n/a	Usage of paws, mouth, dropping pasta, head position Duration of eating Number of pauses	Scores 0-1 for pasta holding and eating with contralesional limb Duration in seconds	Allred <i>et al.</i> , 2008 d'Avila <i>et al.</i> , 2012
Grid test, Ladder rung walking task	TBI (8 w.) Stroke (6 w.)	n/a	Number of foot faults Number of slips Qualitative analysis of foot faults	Foot faults with both forelimbs Number of foot faults relative to walking time Qualitative foot fault score (0-6)	Ajao <i>et al.</i> , 2012 Farr <i>et al.</i> , 2006 Goffus <i>et al.</i> , 2010 Onyszchuk <i>et al.</i> , 2007

Cylinder test	TBI (6 w.) Stroke (17 w.)	n/a	Wall contacts with individual forelimbs and both forelimbs Paw used for landing	Contralesional forelimb contacts relative to all contacts Asymmetry scores (comparison of contralesional and ipsilesional forelimb)	Metz <i>et al.</i> , 2005 Schallert <i>et al.</i> , 2000 Schönfeld <i>et al.</i> , 2016
Grip strength test	TBI (4 w.) Stroke (2 w.)	n/a	Force in grams	Grams for both forelimbs or hindlimbs	Evans <i>et al.</i> , 2015 Madinier <i>et al.</i> , 2014

TESTS FOR GAIT AND INTER-LIMB COORDINATION					
Test	Animal models	Behavioral assessment	Readout	Data presentation	References
Basso, Beattie and Bresnahan (BBB) scale	SCI (rats; 6 w.)	21-point rating scale	Joint movement Limb placement	Score between 0-21	Basso <i>et al.</i> , 1996 Inoue <i>et al.</i> , 2013
Basso mouse scale (BMS)	SCI (mice; 6 w.)	9-point rating scale	Paw position Limb placement	Score between 0-9	Basso <i>et al.</i> , 2006
CatWalk XT	TBI (5 w.) Stroke (6 w.) SCI (12 w.) Parkinson's disease (4 w.) Muscle denervation (8 w.)	n/a	Gait-related parameters	Scores for each paw separately Scores for all four paws	Chen <i>et al.</i> , 2015 Mondello <i>et al.</i> , 2015 Mountney <i>et al.</i> , 2013 Schönfeld <i>et al.</i> , 2016 Wang <i>et al.</i> , 2011 Westin <i>et al.</i> , 2012

Rotarod	TBI (4 w.) Stroke (4 w.) SCI (1-2 w.)	Accelerated speed Constant speed	Latency to fall off the rod	Latency in seconds	Brody <i>et al.</i> , 2007 Kim <i>et al.</i> , 2015 Onyszchuk <i>et al.</i> , 2007 Vidal <i>et al.</i> , 2013 Vonder Haar <i>et al.</i> , 2012
Beam walking	TBI (< 1 w.) Stroke (4 w.)	With or without supporting ledge	7 point scale to evaluate walking Time to cross beam	Score between 0-7 Time in seconds	Qu <i>et al.</i> , 2015 Shijo <i>et al.</i> , 2015

NEUROLOGICAL DEFICIT SCORES					
Test	Animal models	Behavioral assessment	Readout	Data presentation	
Bederson scale	Stroke (> 1 w.) TBI (4 w.)	n/a	Paw asymmetry and balance, general mobility	3 tests, 0-3 points each	Bederson <i>et al.</i> , 1986 Mountney <i>et al.</i> , 2013
Garcia scale	Stroke (> 1 w.)	n/a	Paw asymmetry, general mobility, grip strength, reaction to sensory stimuli	6 tests, 0-3 points each	Garcia <i>et al.</i> , 1995
Modified neurological severity score (mNSS)	TBI (5 w.) Stroke (2 w.)	n/a	Paw asymmetry and balance, general mobility, reaction to sensory stimuli	5 tests, 0-6 points each	Chen <i>et al.</i> , 2001 Wang <i>et al.</i> , 2013

Table 2: Six recommendations are shown that should assist in the selection of the best behavioral test. BBB: Basso, Beattie and Bresnahan rating scale; BMS: Basso mouse scale; SCI: spinal cord injury.

ADVICE	THIS MEANS	EXAMPLE
Complex tests are better than simple tests	Complex tests measure <i>multiple aspects of one type of behavior</i> , e.g. gait	Complex test: BBB/BMS Simple test: Rotarod
Multiple readouts are better than a single readout	Multiple readouts lead to more than one <i>outcome variable</i>	Multiple readouts: CatWalk XT (<i>mean intensity, print area, stride length</i> , amongst others) Single readout: Grid test (<i>foot faults</i>)
Avoid multiple tests	Every additional test may <i>stress</i> the animals and <i>takes time</i>	Choose one fine motor skill test instead of combining the single pellet reaching test, the staircase test and the adhesive removal test
Automated tests are better than observer-based scoring	Automated tests are <i>objective</i> and do not require two or more blinded observers	Automated test: CatWalk XT Observer-based test: BBB/BMS
Combine an automated gait test with a fine motor skill test	After <i>TBI or stroke</i> , gait may be intact whereas fine motor skills are impaired	CatWalk XT + staircase test Rotarod + adhesive removal test
Choose the 'gold standard test'	'Gold standard tests' are <i>regularly published</i> for a specific animal model	Stroke: Single pellet reaching test SCI: BBB/BMS