



Published in final edited form as:

Behav Brain Res. 2007 February 27; 177(2): 232–241.

Adaptation of a ladder beam walking task to assess locomotor recovery in mice following spinal cord injury

Brian J. Cummings¹, Christie Engesser-Cesar¹, and Aileen J. Anderson^{1,2,3}

¹ *Physical Medicine & Rehabilitation, University of California, Irvine, CA 92697*

² *Anatomy & Neurobiology, University of California, Irvine, CA 92697*

³ *Reeve Irvine Research Center, University of California, Irvine, CA 92697*

Abstract

Locomotor impairments after spinal cord injury (SCI) are often assessed using open-field rating scales. These tasks have the advantage of spanning the range from complete paralysis to normal walking; however, they lack sensitivity at specific levels of recovery. Additionally, most supplemental assessments were developed in rats, not mice. For example, the horizontal ladder beam has been used to measure recovery in the rat after SCI. This parametric task results in a videotaped archival record of the event, is easily administered, and is unambiguously scored. Although a ladder beam apparatus for mice is available, its use in the assessment of recovery in SCI mice is rare, possibly because normative data for uninjured mice and the type of step misplacements injured mice exhibit is lacking. We report the development of a modified ladder beam instrument and scoring system to measure hindlimb recovery in vertebral T9 contusion spinal cord injured mice. The mouse ladder beam allows for the use of standard parametric statistical tests to assess locomotor recovery. Ladder beam performance is consistent across four strains of mice, there are no sex differences, and inter-rater reliability between observers is high. The ladder beam score is proportional to injury severity and can be used to easily separate mice capable of weight-supported stance up to mice with consistent forelimb to hindlimb coordination. Critically, horizontal ladder beam testing discriminates between mice that score identically in terms of stepping frequency in open-field testing.

INTRODUCTION

There are many methods to measure recovery of locomotor function after traumatic brain injury or spinal cord injury (SCI) in rodents [18,22]. Most instruments have been developed to assess rats, not mice, due to the more prevalent usage of rats to model injury and disease. With the recent emphasis on molecular mechanisms associated with trauma, many studies have sought to use transgenic mice to test specific molecular pathways. The need to assess locomotion in mice necessitates either adaptation of rodent instruments or development of new instruments. Open-field locomotion in spinal injured rats can be assessed using the Basso, Beattie and Bresnahan locomotor rating scale, the BBB [1]. This test can assess a wide variety of spinal injuries resulting in impairments ranging from partial joint movement, to weight supported standing, through coordinated walking and trunk stability. Similarly, the Basso Mouse Scale (BMS) was recently developed for open-field locomotor assessment to account for the unique

Corresponding Author: Brian J Cummings, 1254 Gillespie NRF, University of California, Irvine, CA 92697-4540, Tel: 949-824-3254, Fax: 949-824-9728, cummings@uci.edu

Publisher's Disclaimer: 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 citable 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.

recovery pattern seen in mice [3,8]. Both the BBB and BMS were designed to assess gross recovery levels across the full range of recovery, but can be less sensitive at specific levels of recovery, in part due to the ordinal nature of the scale. This issue may be particularly limiting when discriminating between animals that have achieved some degree of coordination between forelimbs and hindlimbs. Other instruments to evaluate forelimb and hindlimb function of SCI rats include grid walking [18,27], rope climbing [33], inclined plane [28], kinematic analysis [12], gait analysis [13,17], measures of ground reaction forces [23], and swimming [15,19].

The horizontal ladder beam was developed to assess both forelimb and hindlimb deficits following sensorimotor cortex injury in rats. On a horizontal ladder, uninjured rats grasp rungs with their forelimbs; consequently, forelimb footslips (grasping a rung and then slipping off) or forelimb misplacements (wrist stepping) following sensorimotor cortex injury or cervical spinal cord injury are easily identified [30]. Typically, the total number of forelimb footslips or misplacements on three passes across ten to twenty rungs is counted. For the hindlimbs, only footslips are scored in the rat; hindlimb misplacements are not scored because uninjured rats step too many different ways with their hindlimbs to allow for reliable scoring [30]. Additionally, rats may learn to adapt to specific deficits given sufficient training on a fixed ladder [20]. Nonetheless, the horizontal ladder beam has several advantages to other instruments, including minimal training of subjects, easy administration, unambiguously scored parametric data, and a videotaped archival record of the event. Recently, the ladder beam has been used extensively to quantify recovery of function in spinal injured rats [4,5, 11,16,19–21,25,26,32].

The horizontal ladder beam is suitable for use with SCI rats that are capable of weight supported stepping. Within this range, the ladder beam has been shown to detect subtle hindlimb differences in “recovered” versus normal rats [20,30,31]. For example, Liebscher and colleagues demonstrated a significant improvement in locomotion following Nogo-A antibody administration after a T-hemisection [19]. Rats that received Nogo-A antibodies showed a 2 point difference on the BBB compared to animals treated with a control IgG (13 vs 15) but greater than a 20 percent difference on the ladder beam. Very similar sized relationships between the performance on the BBB and ladder beam were recently reported in SCI animals treated with hydrogel and trophic factors compared to control animals [26].

With the increasing use of transgenic mice to model neurotrauma and SCI, there is a greater need for sensitive, reliable, quantitative tests of the recovery of locomotor function that facilitate parametric statistical tests between treatment versus control groups. Although a smaller ladder beam apparatus for mice is commercially available, its use in the assessment of recovery in spinal injured mice is rare [6,8]. One other group used a ladder assessment of functional recovery in mice, but mice traversed either an ascending or a descending ladder, not a horizontal ladder [9]. Further, an assessment of the types of misplacements injured mice exhibit and normative data for mice on the horizontal ladder beam task is lacking. Thus, it is unclear whether hindlimb misplacements can be reliably identified in partially paralyzed mice, or whether the types of misplacements are important. As with any new instrument, the ladder beam must demonstrate greater sensitivity and reliability than other instruments, or be easier and quicker to administer, or enable measurement of a new characteristic of recovery compared to measures already in use. We report here the characterization of a modified ladder beam instrument and scoring system to supplement open-field testing. The horizontal ladder beam task quantifies hindlimb recovery in spinal contused mice, is easy to administer, and is more sensitive to recovery of stepping and coordination than open-field testing.

METHODS

Subjects

Four strains of mice were tested to determine if there were inter-strain differences in locomotor performance on the ladder beam task in uninjured mice. Adult NOD-*scid*, Bub/BnJ and Balb/c mice were obtained from Jackson Laboratories. C57Bl/6 mice were obtained from Harlan-Sprague Dawley, Inc. Animals were group housed; five per cage, on corn cob bedding with access to food and water *ad libitum*, and holding rooms were maintained on a 12-hour light/dark cycle. The Institutional Animal Care and Use Committee (IACUC) at the University of California, Irvine, approved all animal procedures.

Surgical Procedures

Animal subjects were between 7 and 11 weeks of age at the time of injury, and ranged in weight from 16–24 g. All animals were anesthetized with 0.5 ml/20 g Avertin by intraperitoneal injection. Spinal cord contusions were induced after laminectomy of the T9 vertebra using the Infinite Horizons (IH) force-controlled Impactor device (PSI, Lexington, KY), which uses a stepping motor and probe to impact the spinal cord and a force sensor to detect the actual force during impact [29]. There is a linear relationship between impact force and tissue sparing ($r=0.67$). Impact forces were set at 30, 50, or 60 kdynes to obtain predetermined levels of injury roughly corresponding to 70%, 50% or 35% tissue sparing, respectively, at the epi-center [10]. The impactor tip was centered medio-laterally and rostro-caudally 4 mm above the spinal cord. Following impact, bruise severity was determined under a microscope; animals with unilateral bruises were excluded. A small piece of absorbable gel foam was used to cover the exposed spinal cord and the overlying muscle was sutured with 5-0 Chromic Gut. The skin incision was closed with 7 mm wound clips. Animals were given 2.5 mg/kg Baytril (Enrofloxacin), 2 ml lactated ringers, and 0.05 mg/kg buprenorphine subcutaneously before being placed in recovery cages.

Fresh recovery cages were prepared with AlphaDri bedding (Newco Distributors, Inc. Rancho Cucamonga, CA), food available *ad libitum* was placed on the bottom of the cage, and water bottles with long sipping tubes were used. Buprenorphine (0.05 mg/kg) was given subcutaneously at 48 hours post-injury. Baytril (2.5 mg/kg) was administered prophylactically by subcutaneous injection once a day for 14–21 days post surgery. Lactated Ringers (2 ml) was given subcutaneously once daily for 3–7 days post-operatively. Bladders were manually expressed twice daily for approximately 21 days or until animals had regained partial voluntary micturation, at which time they were reduced to once a day manual expression.

Open-Field Behavioral Assessment

An additional set of animals collected from different studies in our lab ($n=125$) were assessed on a modified open-field BBB scale (0–18) by two experienced raters prior to ladder beam assessment to allow for comparisons on the ladder beam task to a more commonly used assessment tool. At the time these animals were assessed behaviorally, the Basso Mouse Scale was still under development, thus BMS scores were not available. Because toe clearance is difficult to assess in mice due to their small size, Dergham and colleagues modified the BBB to eliminate categories based on toe clearance (rat scores 16, 17, 18) [7]. They also collapsed the remaining two categories (consistent stepping and coordination, parallel paw position, trunk instability/stability, and consistent tail up) for a 0–17 point scale. In our hands, trunk instability (17) and trunk stability (18) are discretely observable and were included in our modified BBB scale.

Ladder Beam Apparatus

The apparatus consists of a metal horizontal ladder beam with 74 rungs suspended 18 inches above the ground with a hollow black escape box at one end. The commercially available ladder for mice has 4 mm diameter rungs spaced 12 mm apart (Columbus Instruments, Ohio) and can be outfitted to automatically detect foot faults. We found the automated system inaccurate and insensitive for mouse footslips and misplacements, necessitating videotaping. A Canon Elura 20MC progressive scan digital video camcorder was used to film the trials in progressive scan mode (odd and even lines are scanned each frame which effectively doubles the video resolution for frame by frame analysis). Four clip-on fluorescent lamps (60W) were used to illuminate the rungs from underneath the ladder, two on the start side and two on the end side of the support apparatus. Auto white balance and auto shutter modes were sufficient with this level of illumination. A custom-made Plexiglas sliding platform below the ladder beam supported the digital video camera to allow tracking of the animal while crossing the ladder. Individual rungs were numbered along one side and every tenth rung was identified by different colored tape. The first rung scored, labeled rung 1, was 10 cm from the actual start of the ladder so that animals would be moving at a constant rate prior to being scored. Rungs were numbered from 1 to 50, and rung numbers and colored tape at every tenth rung were visible through the camera to facilitate analysis. A three-sided tunnel constructed of foam-board, 8 cm wide and 10 cm tall, was placed on the ladder beam for each animal to run through towards the escape black box to keep the subject within the camera's field of view during taping and reduce the tendency for the animal to reverse direction mid-run. Bright, even illumination and a progressive scanning mini-DV video camera are essential to obtain high quality individual video frames suitable for analysis.

Videotaping

Animals were pre-handled for one week before their first videotaping and trained on the ladder beam apparatus three days prior to assessment. Training consisted of a 5-minute acclimation period where the animal was placed in the escape box (containing bedding from their home cage, food and a paper towel), followed by at least three trials where the animal was directed to run across the ladder beam towards the escape box. In between trials, each animal was allowed to remain in the escape box for one minute to acclimate. Prior to taping trials for an individual animal, a white dry erase board with the date and animal ID number was briefly recorded, the camera "paused" and then the first trial recorded. Each animal was filmed for a minimum of four good trials. A good trial consisted of filming the hindlimbs of the animal while crossing at a constant rate across all 50 rungs without turning around. Between trials, the camera was again "paused", and the animal left in the escape box for one minute. For injured animals, pre-training was performed again a few minutes prior to each animal's taping by re-introducing the animal to the escape box.

Ladder Beam Video Analysis

Digital videotapes were transferred to a Macintosh computer running version 5.0 of iMovie using a firewire cable. This combination of DV video camera and software enabled the automatic transfer of video files with a break inserted between each clip based on when the video camera was "paused" during recording. Thus, for each animal, trials appear sequentially as numbered video clips, separated by a white clip with the animal ID number of the next animal. The clip number and animal number were recorded in a spreadsheet with the rater's assessment of stepping for each rung (see Table 1). As a rule, the first three assessable clips were used and extra clips discarded. Assessable clips were defined by sharp, focused video and movement at a constant rate, without the animal pausing or reversing direction between rungs 1 through 50.

Two independent observers, blind to both the animal's group data and each other's assessment, scored 3 passes per animal in an Excel spreadsheet with data integrity checking and auto-scoring. Post-hoc analysis of injured and normal mice enabled hindlimb stepping to be binned into discrete, frequently occurring categories. We observed three common "positive" stepping events and a variety of foot-faults which could be condensed to three "negative" categories. Plantar grasping of the rung was the most common type of rung stepping observed in uninjured animals: "Plantar"; followed by contact with the rung by the toes only: "Toe". A "Skip" is scored when the animal's hindlimb passes over a rung during a step cycle. Since Skips are part of a normal stepping on a horizontal ladder and occur with high frequency (~70%), we consider Skips a positive event. A "Miss" was scored when the hindlimb drops below the plane of the rungs without touching either the lower numbered rung or higher numbered rung. The "Miss" score was assigned to the lower numbered rung and the next rung was scored fresh based on the next attempted step cycle. A Miss indicates that a complete step cycle was attempted, but no rung was contacted, as opposed to a Slip, where there was rung contact, or a Drag, where a step cycle was not completed. The Drag category indicates either a true "drag" or incomplete step cycle, or any other unclassifiable event (e.g. spasm during a step).

Based on the literature and performance of SCI injured rats, we initially included separate negative categories for undefined/doubtful events (U) and spasm (SP), as well as additional positive categories {e.g. ankle (A) and dorsal (D) stepping} that were subsequently dropped after scoring numerous injured and non-injured animals. We found that these events happened too rarely in mice to warrant separate categories. Dropping these rare, extraneous categories did not reduce the ability to separate animals based on injury severity, while greatly decreasing the time required to score an animal and increasing the inter-rater reliability (data not shown). As with scoring animals in the open field on the BBB/BMS, when in doubt, score to the deficit to increase one's ability to detect changes over time. This approach is intentionally biased against rejecting the null hypothesis. Using the operational definitions contained in Table 2, events were scored and entered into a spreadsheet as a single letter for each rung.

Although it is difficult to illustrate movement with still images, by pseudo-coloring a fixed bar position in each series of images, the six discrete events can be seen across sequential video frames (Figure 1). In each example, the colored bar is the bar(s) being scored. Results can be reported as either a Ladder Beam Score % (LBS %: the number of positive events/total number of events) that allows for comparisons between animals that did not finish three complete passes or a Cumulative Error score (CE: the total number of cumulative errors over three ladder passes).

Statistical Approach

Four strains of mice were tested for their ability to traverse a horizontal ladder beam when uninjured. For one strain (Bub/BnJ), both males and females were tested (independent variable). After normative data in uninjured mice were obtained, mice received a spinal cord injury at one of three different severities and were tested again. The dependent measures in this study were 6 possible event categories for stepping on a ladder rung, a Ladder Beam Score computed as a percentage from these 6 event categories, and a Cumulative Error score (the sum of the 3 negative event categories). Variables were tested for normalcy by histogram plots and homogeneity of variance by Bartlett's test. A one-way Model 1 ANOVA (fixed treatments) was used to test for significant differences between the means of the various independent measures (strain, gender, injury severity) and LBS, CE, or individual events on the ladder beam. Where unequal variance was observed (the 3 negative event categories), a Welch ANOVA for the means was to correct for unequal variance and provide a corrected F statistic. The effect of gender on LBS was tested using an unpaired, two-tailed student's t-test. To test for the discriminative power of LBS to predict open-field performance one week later on the BBB, a

Chi square test for a trend across three levels of LBS was performed in Prism 4.0. All other statistical analyses were done using JMP 6.0.

RESULTS

Inter-rater reliability

A comparison of the ladder beam score (LBS) assigned to an animal by observer one compared to the score assigned by observer two yields an inter-rater reliability (IRR) for those two observers. Observers are trained with a set of four test animals (each with three passes across the ladder) ranging from mildly impaired to severely impaired. Observers must achieve 95% agreement with a concealed “answer key” to become certified to score animals independently. A typical undergraduate student can reach this level of proficiency after scoring these test animals during 8 to 12 hours of practice. Once observers have achieved competence on the test animals, a recheck of inter-rater reliability among four trained observers on actual experimental animals (n=18) yielded a Pearson’s IRR = 0.948 on the Ladder Beam Score.

Comparison of performance across non-injured mice strains and sex

Four strains of mice, NOD-*scid* (n=10), BuB/BnJ (n=20), Balb/C (n=12), and C57Bl/6 (n=10), were tested to determine if there were strain differences in locomotor performance on the horizontal ladder beam task. Uninjured, naïve mice were scored by two observers on the six events outlined above (Table 2) and a percent LBS was computed. Uninjured mice performed the ladder beam task with almost no errors; the average percent LBS was 99.07 (SE ± 0.14) (Figure 2A). Unequal variance on the LBS in Bub/BnJ mice was observed, necessitating use of a Welch ANOVA to obtain a corrected F value when comparing strains. There were no significant differences between Balb/C, C57BL6 and Nod-*scid* mice (strain as the independent variable) on a one-way Welch ANOVA by percent LBS (dependent variable) $F(3,48) = 14.49$, $p \leq 0.0001$. Bub/BnJ mice made less than one more error over three passes than the other strains (Tukey-Kramer, $q=2.66$, $\alpha=0.05$) [Abs(Dif)-LSD of Bub/BnJ vs C57BL6 = 0.84, vs Nod-*scid* = 0.79, and vs Balb/C = 0.54]. Technicians noted that Bub/BnJ mice were more agitated during pre-training and acclimatization than the other strains, which may account for their slightly lower LBS. Although the large number of Bub/BnJ in this analysis resulted in a statistically significant difference, it is unlikely that this one point variation is biologically significant in terms of sensitivity or applicability of this task to analysis in a research setting. Excluding Bub/BnJ mice, the average percent LBS for the other three strains of uninjured mice is 99.63 (SE ± 0.09). While most spinal cord injury models use female mice, it is also important to know if there are differences in performance on the ladder beam task in female versus male mice. A unpaired Student’s t-test of 20 non-injured Bub/BnJ mice (n=10 females) shows no gender differences in the performance of normal mice on the task [Abs(Dif)-LBS = -0.27 male vs. female, $p \leq 0.05$] (Figure 2B). The number of individual events (plantar steps, skips, etc.) did not show gender differences either (data not shown).

Relationship between injury severity and ladder beam performance

For the ladder beam task to have external validity, animals should perform worse on the task with increasing severity of spinal cord injury. We compared percent LBS and cumulative errors (CE) in mice that received no injury (n=32), versus a mild (30 kD, n=13), moderate (50 kD, n=24), or more severe (60 kD, n=34) contusion injury. The percent LBS decreased with increasing injury severity ($r = -0.72$, $p < 0.001$) while CE increased with injury severity ($r = 0.72$, $p < 0.001$). A one-way ANOVA of injury severity (independent variable) by either LBS or CE (dependent variable) indicates significant differences in performance by injury severity level [$F(3,72)=77.12$, $p < 0.0001$ for LBS; 77.30 , $p < 0.0001$ for CE]. For both measures, there was a significant difference between both the 50 kD and 60 kD groups compared to uninjured animals using Dunnett’s method of assigning the uninjured animals as the fixed “control” group

(Dunnett's $|d|$ 3.76 $p < 0.001$). The 30 kD group was not detectably different from uninjured controls on either the percent LBS or CE, but was different from both 50 kD and 60 kD groups (Student's t -test, $t=3.43$, $p<0.001$) (Figure 3). Results were similar when individual positive or negative events were examined separately. Both 50 kD and 60 kD groups were significantly different from uninjured controls on plantar, toe, skip, and miss events; only the 50 kD group was significantly different from controls on slips; while only the 60 kD group was different from controls on drags (Dunnett's $|d|$ 3.02 $p < 0.01$) (Figure 4A and 4B).

Comparison of ladder beam performance to open-field performance

Open-field rating scales, while extremely useful as a general screen in assessment of locomotor recovery after SCI, are limited in their ability to discriminate some areas of recovery of function by the ordinal nature of these scales. For example, an animal that successfully plantar steps once or achieves a single coordinated bout of locomotion during 4 minutes of forward motion receives the identical score as an animal that successfully plantar steps or achieves coordinated locomotion 49% of the time ('occasional' stepping on the BBB, 'some' stepping on the BMS). Yet these two animals are not, in fact, equally recovered. For external validity, one would predict that there is a relationship between performance in the open-field and on the ladder beam task for any given animal tested on both tasks within 24-hours. Although the ladder beam task is linear and continuous, the BBB is not; thus, one would not expect the relationship between the BBB and LBS to be linear. Indeed, the best-fit model between performances on both tasks across a broad range of animals ($n=125$) is a third order polynomial. If $BBB=y$ and $LBS=x$, then $y=6E5x^3 - 0.0089x^2 + 0.3885x + 5.1825$ ($r^2=0.8197$, $r=0.91$). Plotting the non-linear relationship demonstrates that there is a discrete window of locomotion where the ladder beam may be more sensitive than open-field testing, and that the LBS would be most advantageous for separating animals with occasional stepping through frequent forelimb-hindlimb coordination (Figure 5).

Discriminative power of LBS on subsequent locomotor performance

We hypothesized that a high percent LBS should predict improved performance on the BBB over time, particularly in the range on the BBB where animals with large differences in ability could potentially receive similar scores. Conversely, a low percent LBS would be more likely in animals that barely reached criterion on the BBB, hence their 1 week follow-up BBB would be more likely to decrease. Thus we selected all animals with an initial BBB score between 9 and 13, indicating that they were in the range of weight supported stance up to stepping with frequent to consistent plantar placement. Next, we excluded animals with no follow-up BBB test one-week post ladder beam assessment. Our dataset included 34 animals. Referencing the BBB to percent LBS curve (Figure 5), we assigned these 34 animals into those with poor ladder beam performance ($LBS \leq 70$, $n=7$), intermediate performance ($71 \leq LBS \leq 89$, $n=13$), and high performance ($LBS \geq 90$, $n=14$). This analysis reveals that the percent LBS helps predict whether a given animal will maintain or improve their performance on the BBB one week later or get worse. Less than half of animals with a percent LBS below 70 improved on the BBB one week later compared to 69% of animals in the intermediate LBS group and 86% of animals in the high LBS group. A Chi square test for a trend across the three levels of LBS indicates a significant trend (Chi-Square = 4.06 (df 1), $p=0.04$). This analysis demonstrates that percent LBS is capable of differentiating animals that receive the same open-field rating and predicting their future improvement in the open-field (Figure 6).

Example of the detection of differences in hCNS-SC treated animals vs. controls

Three additional points about the horizontal ladder beam task are illustrated by Figure 7. The first is that the ladder beam task can also be used to measure the cumulative number of errors (CE) rather than a percentage of positive events. In cases where the number of errors is small,

this may better illustrate group differences. Second, Figure 7 demonstrates a published example where the measured differences in ladder beam errors between groups (SCI injured mice receiving either vehicle injections or neural stem cell grafts) closely tracks performance assessed in the open-field [6]. Finally, Figure 7 demonstrates the discriminative ability of the ladder beam task in comparison to the BBB. Vehicle versus human neural stem cell grafted animals (hCNS-SC) injured with 50kD T9 contusion were significantly different on both the BBB and LBS, while animals that received a more severe 60kD contusion were not different from controls when assessed on the BBB. However, significant group differences in recovery in the 60kD experiment were detected by the cumulative number of ladder beam errors, where hCNS-SC treated animals made half the number of errors that control animals made [6].

DISCUSSION

A number of instruments are available to assess locomotor recovery following spinal cord injury. Due to the range of possible outcomes following SCI, from complete hindlimb paralysis to normal locomotion with slightly impaired trunk stability or paw position, a single instrument cannot differentiate with equal sensitivity across such a broad spectrum of recovery. Perhaps the best broad-range instrument for assessing locomotor recovery in rats is the BBB [1,2]. Recently, an open-field locomotor rating scale specialized for mice has been developed, the Basso Mouse Scale (BMS) [3]. The BMS was under development during most of the experiments presented in this study, thus BMS scores were not available. We do not recommend using the modified mouse BBB outlined here in future mouse studies; BMS is more appropriate. Open-field rating scales, while extremely useful as a general screen of locomotor recovery after SCI, are limited in their ability to discriminate some areas of recovery of function by the ordinal nature of the scale. For example, an animal that successfully plantar steps once or achieves a single coordinated bout of locomotion during 4 minutes of forward motion receives the identical score as an animal that successfully plantar steps or achieves coordinated locomotion 49% of the time ('occasional' stepping on the BBB, 'some' stepping on the BMS). Nonetheless, both open-field instruments were designed to quantify hindlimb locomotor recovery and coordination across the full range of impairment. The authors of these open-field rank-ordered rating scales recommend that supplemental tasks be used to more finely evaluate animals that fall within particular ranges on either scale. For example, in mice, it is particularly difficult to reliably differentiate between animals within the range of forelimb to hindlimb coordination [3].

The present study demonstrates that the horizontal ladder beam can be reliably administered and scored by different raters with high inter-rater reliability, and that the percent LBS is particularly useful for discriminating between discrete levels of recovery once mice have achieved weight-supported stepping. The LBS is most advantageous for separating animals within the 5 to 7 range on the BMS or the 9 to 12 point range on the "mouse" BBB {corresponding to animals with weight supported stance (9), occasional stepping (10), frequent to consistent plantar stepping with no coordination (11), occasional forelimb-hindlimb coordination (12), to frequent forelimb-hindlimb coordination (13)}. This is a range where animals with large differences in functional recovery can often receive similar or identical open-field scores.

Performance on the horizontal ladder beam corresponded well to the actual force recorded for each injury, demonstrating external validity. When used in mice, the IH device has been shown to produce graded injuries in relation to force [10,24]. As predicted, the percent LBS decreased linearly with 30, 50, and 60 kD force contusion injuries at the T9 level in mice. Notably, the 30 kD group did not differ significantly from uninjured controls on either the percent LBS or CE, suggesting a ceiling effect on this task. This is likely related to locomotor outcome; with this mild of an injury, animals are not in the range of function where we would predict the LBS

to assist in discriminating outcome (*e.g.* this task is insensitive to paw position). However, the 30 kD group was statistically different from both 50 kD and 60 kD groups. The results were similar when individual positive or negative events were examined separately, suggesting that the tracking of discrete types of successful steps or errors is not critical; rather an aggregate score is sufficient. Additionally, the task might not be challenging enough for animals with mild injury, necessitating the use of an inclined ladder, or a ladder with random, variable rung spacing.

There were no significant differences in ladder beam performance between three strains of mice assessed on the ladder beam task (C57B6, Balb/C and Nod-scid, average LBS=99.6). C57BL/6 and Balb/C mice were chosen based on the strains most commonly used in the SCI field. NOD-scid mice are the model of choice for xenograft experiments; a recent trend in the SCI field has been to test various stem cell therapies in these animals. Bub/BnJ mice were chosen because their complement sufficient immune system most closely models the human immune system and inflammation plays a large role in the sequela of spinal injury. Uninjured Bub/BnJ mice scored 1 point lower than the other three strains. Interestingly, differences in locomotion between strains based on coat color have been reported, with black strains exhibiting a narrower base of support compared to brown (F1) or white (Balb) strains [3]. Technicians noted that Bub/BnJ mice were clearly more agitated when pre-handled compared to other strains and that Bub/BnJ mice did not move as willingly along the ladder beam during taping. This could account for the slightly lower average percent LBS of Bub/BnJ mice compared to the other strains. Pre-training or training coupled with a reward might improve Bub/BnJ pre-injury performance, but this is not necessary if all animals in a study are of the same strain. Furthermore, it is unlikely that the variation between Bub/BnJ mice and the other three strains is biologically significant in terms of sensitivity or applicability of this task in a research setting. The majority of SCI studies use female rats or mice; some studies have used males to model specific aspects of SCI. We found no gender differences on the ladder beam task, either for the overall LBS, or on any of the specific event categories. If a new strain of mice is to be assessed on the ladder beam task, it is recommended that naïve and graded injuries be tested first to establish a baseline for the new strain in relation to C57B6 mice.

Mice differ from rats on ladder beam because the types of events observed after injury are different. In rats, forelimb misplacements are quantifiable (as wrist stepping rather than plantar stepping) during recovery from cortical injury, but hindlimb misplacements are not assessed. This is because hindlimb misplacements are observable in uninjured rats, and hindlimb step methods vary. Thus, only hindlimb footslips are quantified [30]. In the present study, less than 1 misplacement per 100 steps occurred in three of four strains of uninjured mice, making the observation of hindlimb slips or misses in mice a relevant event after SCI. Nonetheless, our data in mice suggest that it is not necessary to score each positive and negative event subtype; rather, one can simply score stepping at each rung as a positive event (comprising either a plantar or toe step, or a skip) or a misplacement (including a slip, miss or drag). Neither the resulting LBS nor the CE would change, yet assessment would be easier. Indeed, even when the overall CE for a treatment group is not statistically different from a control group, individual measures of sprouting/contacts a BDNF treated group correlated with CE on a ladder beam task in rats [32]. We suggest that one report either the percent LBS (the number of positive events/the number of all events assessed*100) or the CE (the cumulative number of errors over three trials) to aid inter-laboratory comparisons in the SCI field.

There are several advantages to the use of the ladder beam for assessment of locomotor recovery in mice. Each animal's performance is permanently recorded on videotape and copies are transferred to computer for analysis. Inexpensive tapes can be saved long-term for archival purposes, while the computer files can be analyzed by multiple individuals to ensure the task is being scored accurately and reliably. Another advantage is that the percent LBS is easy to

comprehend. Percent LBS or the number of errors (CE) on the ladder beam is suitable for analysis by standard parametric statistics, rather than requiring the use of less sensitive non-parametric procedures. Unlike rating scales, where a 1-point difference may be highly significant biologically (the difference between foot placement and weight supported stepping) or not (differences in joint movement), interpretation of the LBS is aided by its ordinality. An LBS improvement of 10 is meaningful whether at the low range of the scale or the high range. Finally, the cost of the necessary equipment is relatively low (*e.g.* compared to the CatWalk) and the software to analyze video frame by frame on a Macintosh (iMovie) is free.

There are drawbacks to the use of the ladder beam task in mice. Ladder beam performance is only meaningful in animals capable of weight supported stepping. Furthermore, the LBS cannot differentiate between animals at the highest end of recovery (*e.g.* animals with consistent coordinated walking but varying degrees of trunk instability). Analysis is more time consuming than open-field testing, making the collection of recovery over time more difficult. Repeated assessment on the ladder beam is also compounded by potential learning effects [20], a problem that is avoided by variable rung spacing, or if only a terminal ladder beam assessment is used. This task, as validated here for mice, is designed to be a terminal, supplemental assessment to open-field testing.

A trained observer can score three ladder passes in roughly 30 minutes. Simply scoring each event as positive or negative, however, would speed up the analysis. If one plans to implement such a system, an option would be to increase the rung spacing for mice to 18–20 mm between rungs. This would greatly reduce the number of “double-steps”, a paw landing on two rung simultaneously, and speed up analysis. Columbus Instruments will custom build ladders with specific rung spacing. Finally, when assessing mice, we do not recommend the use of the automated foot fault detection system; it is insensitive to foot slips and thus under-records errors.

There are a wide range of assessment tools to measure recovery of function following spinal cord injury in rats (reviewed here [18,22]). In mice, few assessment methods have been validated. While the ladder beam provides information about an animal’s ability to locomote on discrete rungs, it does not give information about step kinematics, (*e.g.* step width, step length, paw rotation) which can be gained from more labor-intensive analysis such as kinematics or CatWalk assessment [14]. When using regularly spaced rungs and only scoring the hindlimbs, the horizontal ladder beam cannot detect “uncoupling” of the fore- and hindlimbs nor a change in the average stride length [4]. Nonetheless, in mice, the horizontal ladder beam can assist in separating performance in animals at the low end of weight supported walking and coordination, up to consistent coordinated walking. More sensitive tests are necessary to discriminate between animals at higher levels of recovery.

CONCLUSIONS

In mice, the ladder beam task and video analysis to obtain a percent ladder beam score (LBS) or the number of cumulative errors (CE) is useful for assessing mice capable of weight supported stepping and no forelimb to hindlimb coordination up to mice with consistent coordination. While the time required to score performance on the ladder beam task is greater than that to assess mice on either the BBB or BMS, the ladder beam task yields quantitative parametric data that discriminates between mice with different injury severities and/or different treatment groups and allows for the use of more powerful statistics. Hindlimb stepping events can be unambiguously evaluated and scored with high inter-rater reliability, enabling the horizontal ladder beam task to help discriminate finer levels of locomotor recovery than open-field assessment.

Acknowledgements

We thank Gilbert Cadena, Leslie Sheu, and Chris Sontag for technical assistance, the CRF Animal Core at UC Irvine and a grant from the Paralysis Project of America (PPA #9526263 to BJC). Upon publication, ladder beam training materials (training video, answer key, and Excel master scoring spreadsheets) will be provided online.

References

1. Basso DM, Beattie MS, Bresnahan JC. A sensitive and reliable locomotor rating scale for open field testing in rats. *J Neurotrauma* 1995;12(1):1–21. [PubMed: 7783230]
2. Basso DM, Beattie MS, Bresnahan JC, Anderson DK, Faden AI, Gruner JA, Holford TR, Hsu CY, Noble LJ, Nockels R, Perot PL, Salzman SK, Young W. MASCIS evaluation of open field locomotor scores: effects of experience and teamwork on reliability. Multicenter Animal Spinal Cord Injury Study. *J Neurotrauma* 1996;13(7):343–359. [PubMed: 8863191]
3. Basso DM, Fisher LC, Anderson AJ, Jakeman LB, McTigue DM, Popovich PG. Basso Mouse Scale for locomotion detects differences in recovery after spinal cord injury in five common mouse strains. *J Neurotrauma* 2006;23(5):635–659. [PubMed: 16689667]
4. Bolton DA, Tse AD, Ballermann M, Misiaszek JE, Fouad K. Task specific adaptations in rat locomotion: runway versus horizontal ladder. *Behav Brain Res* 2006;168(2):272–279. [PubMed: 16406145]
5. Chan CC, Khodarahmi K, Liu J, Sutherland D, Oschipok LW, Steeves JD, Tetzlaff W. Dose-dependent beneficial and detrimental effects of ROCK inhibitor Y27632 on axonal sprouting and functional recovery after rat spinal cord injury. *Exp Neurol* 2005;196(2):352–364. [PubMed: 16154567]
6. Cummings BJ, Uchida N, Tamaki SJ, Salazar DL, Hooshmand M, Summers R, Gage FH, Anderson AJ. Human neural stem cells differentiate and promote locomotor recovery in spinal cord-injured mice. *Proc Natl Acad Sci U S A* 2005;102(39):14069–14074. [PubMed: 16172374]
7. Dergham P, Ellezam B, Essagian C, Avedissian H, Lubell WD, McKerracher L. Rho signaling pathway targeted to promote spinal cord repair. *J Neurosci* 2002;22(15):6570–6577. [PubMed: 12151536]
8. Engesser-Cesar C, Anderson AJ, Basso DM, Edgerton VR, Cotman CW. Voluntary wheel running improves recovery from a moderate spinal cord injury. *J Neurotrauma* 2005;22(1):157–171. [PubMed: 15665610]
9. Fiore C, Inman DM, Hirose S, Noble LJ, Igarashi T, Compagnone NA. Treatment with the neurosteroid dehydroepiandrosterone promotes recovery of motor behavior after moderate contusive spinal cord injury in the mouse. *J Neurosci Res* 2004;75(3):391–400. [PubMed: 14743452]
10. Fugaccia I, Rabchevsky AG, Zhang P, Lump PP, Main JA, Scheff SW. Characterization of a force based computer controlled spinal cord injury device. *J Neurotrauma* 2001;18:1125.
11. Gensel JC, Tovar CA, Hamers FP, Deibert RJ, Beattie MS, Bresnahan JC. Behavioral and histological characterization of unilateral cervical spinal cord contusion injury in rats. *J Neurotrauma* 2006;23(1):36–54. [PubMed: 16430371]
12. Gimenez y Ribotta M, Orsal D, Feraboli-Lohnherr D, Privat A, Provencher J, Rossignol S. Kinematic analysis of recovered locomotor movements of the hindlimbs in paraplegic rats transplanted with monoaminergic embryonic neurons. *Ann N Y Acad Sci* 1998;860:521–523. [PubMed: 9928352]
13. Hamers FP, Koopmans GC, Joosten EA. CatWalk-assisted gait analysis in the assessment of spinal cord injury. *J Neurotrauma* 2006;23(3–4):537–548. [PubMed: 16629635]
14. Hamers FP, Lankhorst AJ, van Laar TJ, Veldhuis WB, Gispen WH. Automated quantitative gait analysis during overground locomotion in the rat: its application to spinal cord contusion and transection injuries. *J Neurotrauma* 2001;18(2):187–201. [PubMed: 11229711]
15. Kim D, Schallert T, Liu Y, Browarak T, Nayeri N, Tessler A, Fischer I, Murray M. Transplantation of genetically modified fibroblasts expressing BDNF in adult rats with a subtotal hemisection improves specific motor and sensory functions. *Neurorehabil Neural Repair* 2001;15(2):141–150. [PubMed: 11811255]
16. Klapka N, Hermanns S, Straten G, Masannek C, Duis S, Hamers FP, Muller D, Zuschratter W, Muller HW. Suppression of fibrous scarring in spinal cord injury of rat promotes long-distance regeneration of corticospinal tract axons, rescue of primary motoneurons in somatosensory cortex and significant functional recovery. *Eur J Neurosci* 2005;22(12):3047–3058. [PubMed: 16367771]

17. Kloos AD, Fisher LC, Detloff MR, Hassenzuhl DL, Basso DM. Stepwise motor and all-or-none sensory recovery is associated with nonlinear sparing after incremental spinal cord injury in rats. *Exp Neurol* 2005;191(2):251–265. [PubMed: 15649480]
18. Kunkel-Bagden E, Dai HN, Bregman BS. Methods to assess the development and recovery of locomotor function after spinal cord injury in rats. *Exp Neurol* 1993;119(2):153–164. [PubMed: 8432357]
19. Liebscher T, Schnell L, Schnell D, Scholl J, Schneider R, Gullo M, Fouad K, Mir A, Rausch M, Kindler D, Hamers FP, Schwab ME. Nogo-A antibody improves regeneration and locomotion of spinal cord-injured rats. *Ann Neurol* 2005;58(5):706–719. [PubMed: 16173073]
20. Metz GA, Whishaw IQ. Cortical and subcortical lesions impair skilled walking in the ladder rung walking test: a new task to evaluate fore- and hindlimb stepping, placing, and co-ordination. *J Neurosci Methods* 2002;115(2):169–179. [PubMed: 11992668]
21. Miya D, Giszter S, Mori F, Adipudi V, Tessler A, Murray M. Fetal transplants alter the development of function after spinal cord transection in newborn rats. *J Neurosci* 1997;17(12):4856–4872. [PubMed: 9169544]
22. Muir GD, Webb AA. Mini-review: assessment of behavioural recovery following spinal cord injury in rats. *Eur J Neurosci* 2000;12(9):3079–3086. [PubMed: 10998091]
23. Muir GD, Whishaw IQ. Complete locomotor recovery following corticospinal tract lesions: measurement of ground reaction forces during overground locomotion in rats. *Behav Brain Res* 1999;103(1):45–53. [PubMed: 10475163]
24. Nishi RA, Liu H, Chu Y, Hamamura M, Su M, Nalcioglu O, Anderson AJ. Measurement of lesion volume by MRI and histological analysis in C57Bl/6 Mice with various severities of T9 spinal cord contusion injuries. 2006submitted
25. Norrie BA, Nevett-Duchcherer JM, Gorassini MA. Reduced functional recovery by delaying motor training after spinal cord injury. *J Neurophysiol* 2005;94(1):255–264. [PubMed: 15985696]
26. Piantino J, Burdick JA, Goldberg D, Langer R, Benowitz LI. An injectable, biodegradable hydrogel for trophic factor delivery enhances axonal rewiring and improves performance after spinal cord injury. *Exp Neurol*. 2006
27. Prakriya M, McCabe PM, Holets VR. A computerized grid walking system for evaluating the accuracy of locomotion in rats. *J Neurosci Methods* 1993;48(1–2):15–25. [PubMed: 8377518]
28. Rivlin AS, Tator CH. Objective clinical assessment of motor function after experimental spinal cord injury in the rat. *J Neurosurg* 1977;47(4):577–581. [PubMed: 903810]
29. Scheff SW, Rabchevsky AG, Fugaccia I, Main JA, Lumpp JE Jr. Experimental modeling of spinal cord injury: characterization of a force-defined injury device. *J Neurotrauma* 2003;20(2):179–193. [PubMed: 12675971]
30. Soblosky JS, Colgin LL, Chorney-Lane D, Davidson JF, Carey ME. Ladder beam and camera video recording system for evaluating forelimb and hindlimb deficits after sensorimotor cortex injury in rats. *J Neurosci Methods* 1997;78(1–2):75–83. [PubMed: 9497003]
31. Soblosky JS, Song JH, Dinh DH. Graded unilateral cervical spinal cord injury in the rat: evaluation of forelimb recovery and histological effects. *Behav Brain Res* 2001;119(1):1–13. [PubMed: 11164520]
32. Vavrek R, Girgis J, Tetzlaff W, Hiebert GW, Fouad K. BDNF promotes connections of corticospinal neurons onto spared descending interneurons in spinal cord injured rats. *Brain* 2006;129(Pt 6):1534–1545. [PubMed: 16632552]
33. Z'Graggen WJ, Metz GA, Kartje GL, Thallmair M, Schwab ME. Functional recovery and enhanced corticofugal plasticity after unilateral pyramidal tract lesion and blockade of myelin-associated neurite growth inhibitors in adult rats. *J Neurosci* 1998;18(12):4744–4757. [PubMed: 9614248]

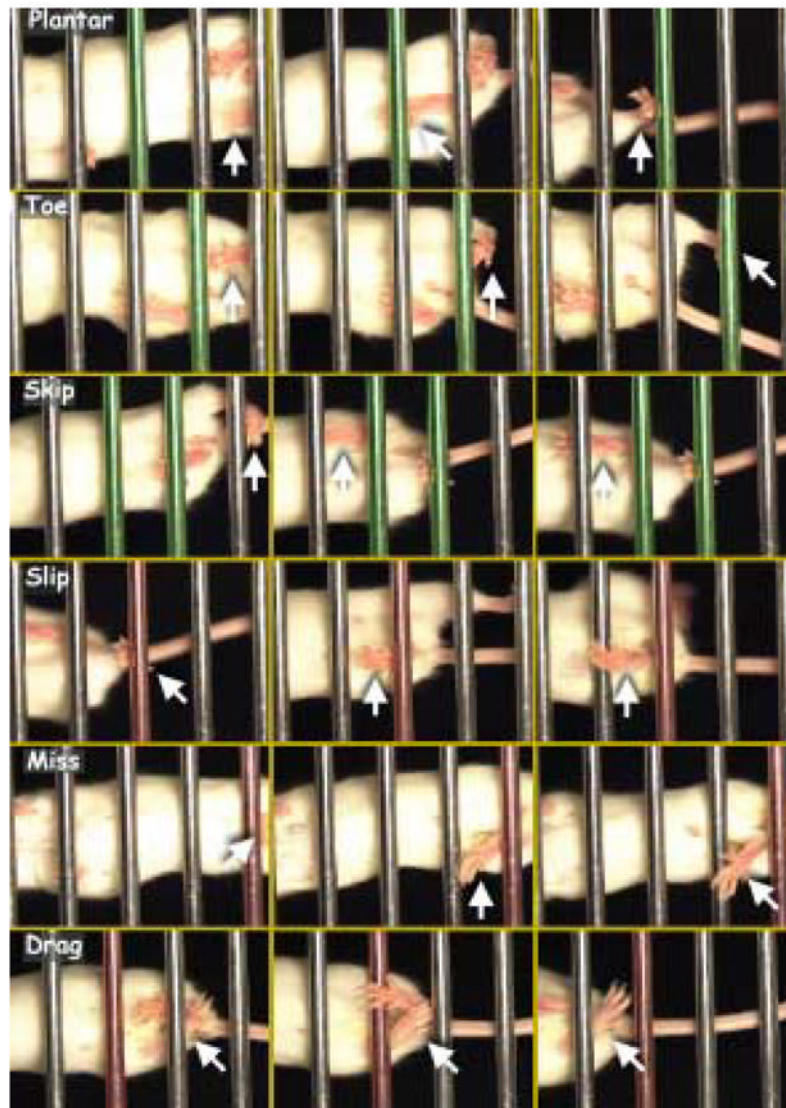


Figure 1. Scoring guidelines for paw placement on the mouse ladder beam

For each row, the pseudo-colored bar is the one being scored. **Plantar**. In frame one, the arrow points to the right hind paw as it approaches the green rung; several video frames later, the right hind paw has made contact (arrow) with the green rung and the toes are visible “above” (in the direction of travel) the rung. In the last frame, the left hind paw is well past the green rung while the right hind paw has almost completed contact (arrow) with the green rung. **Toe**. The arrow points to the left hind paw as it makes initial contact with the green rung. In the next frame, the left hind limb is halfway through a step cycle, yet no toes are visible “above” the rung. In the last frame, there are still no toes visible as the left hind paw leaves this rung.

Skip. In this example, both green rungs are scored as skips because the left hind paw completes a step cycle (starting at the first arrow) and lands on the fourth rung (arrow in last frame). Skips over 3 contiguous rungs are also frequently observed. **Slip**. In the first frame, it appears that the red rung should be scored as a plantar step (arrow). However, in the subsequent two frames, it is clear that the right rear paw slipped off the red rung and dropped below the horizontal plane of the ladder. A slip is differentiated from a miss by paw contact with the rung being scored. **Miss**. Although a rare event, the left hind paw in the first frame does not touch the red

rung, (arrow), but drops below the horizontal plane of the rungs between the red rung and the next rung. The initial red rung is scored as a miss, while the subsequent rung is scored independently. **Drag.** The dorsal surface of the right hind paw contacts the rung preceding the red rung, and in subsequent frames, it can be seen that the dorsal surface drags across the red rung.

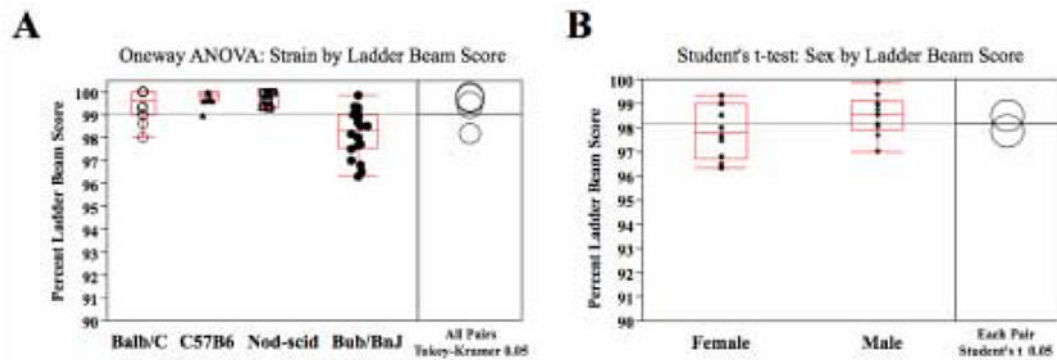


Figure 2. Uninjured Ladder Beam Scores by mouse strain

(A) A box plot shows the 25th to 75th quantiles of LBS per group. There were no significant strain differences between Balb/C, C57BL6, and Nod-*scid* mice on the ladder beam score. However, naïve Bub/BnJ mice averaged approximately one more error compared to the other strains (Tukey-Kramer, $\alpha=0.05$). Excluding Bub/BnJ mice, the average uninjured ladder beam score across strains is 99.63 ($SE \pm 0.09$). (B) There were no significant gender differences between male and female Bub/BnJ mice on the ladder beam score (unpaired Student's t-test, $Abs(Dif) - LSD = -0.27$ male vs. female).

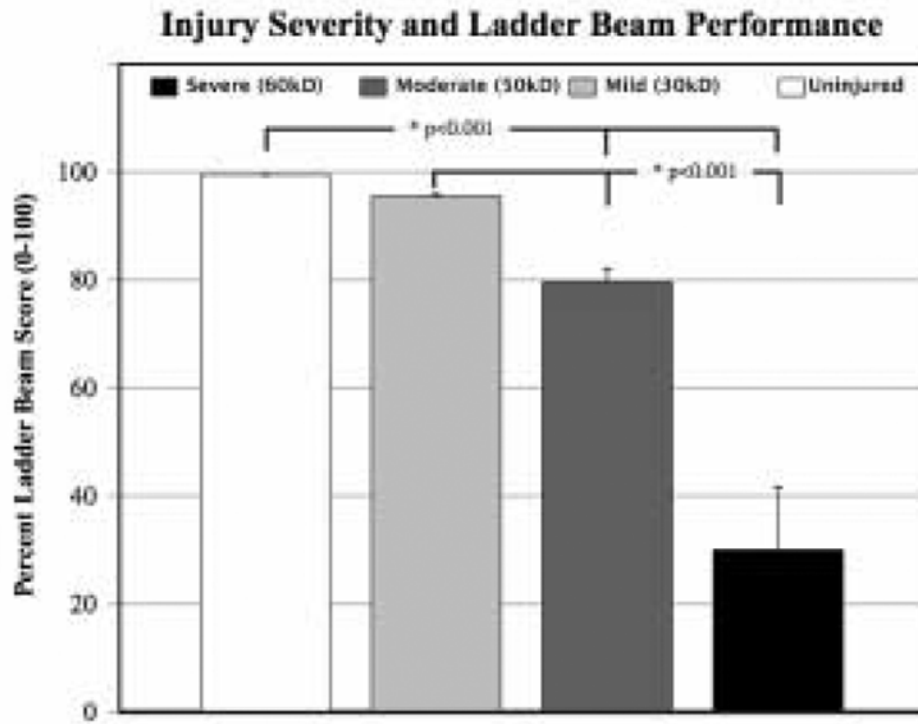


Figure 3. Ladder beam score (LBS) varies inversely in relation to injury severity
 Performance on the horizontal ladder beam was impaired in relationship to the level of force used to generate an injury. Both the moderate (50 kD) and more severe (60 kD) injury groups differed from un-injured controls on the percent LBS and CE (Dunnett's $|d|$ 3.76 $p < 0.001$). The 30 kD group did not differ measurably from controls, but was statistically different from both the 50 and 60 kD groups (Student's T-test, $p < 0.001$). Error bars are \pm SEM.

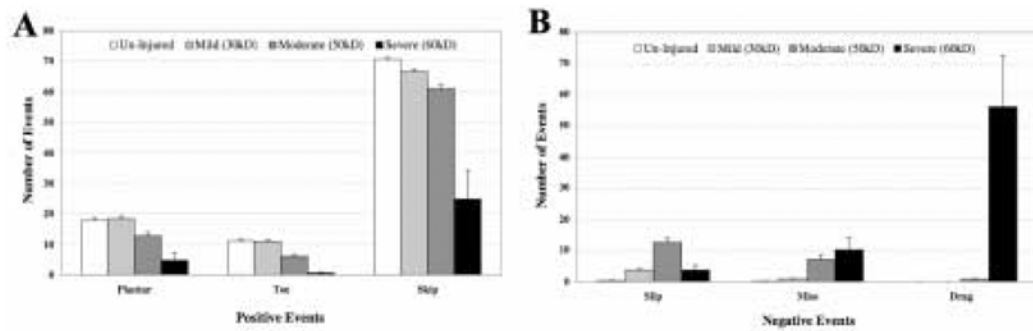


Figure 4. Relationship of positive events versus negative events to injury severity

Six possible outcomes were evaluated for each rung and normalized to 100 possible steps (50/hindlimb). Normal, uninjured mice skip over 70% of the rungs, making plantar steps or toe steps on the remaining rungs. **(A)** The number of individual positive events decreases with increasing injury severity. **(B)** Negative events are extremely rare in uninjured mice (white bars), while the number of negative events increases with increasing injury severity. In the most severely injured mice (60 kD), ~55% of the events observed were “drags”, indicating that these mice were too injured to perform meaningfully on the task. Error bars are \pm SEM. Statistics are reported in the text.

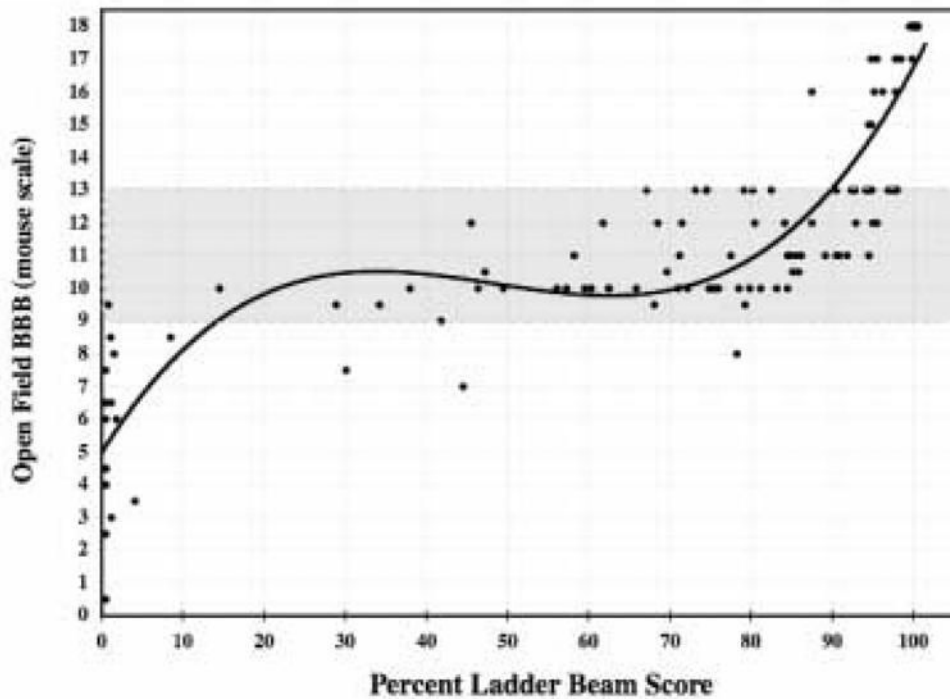


Figure 5. Relationship of percent ladder beam score (LBS) to performance on the “mouse” BBB
 There is a third order polynomial relationship between performance on the BBB and ladder beam score measured within 1 week of each other. If BBB=y and LBS=x, then $y=6E5x^3 - 0.0089x^2 + 0.3885x + 5.1825$ ($r^2=0.8197$, $r=0.91$). This graphical representation suggests that the LBS is optimal to separate animals within the 9 to 13 point range on the mouse modified BBB. The grey box corresponds to animals with weight-supported stance up to animals with frequent to consistent plantar stepping with frequent forelimb to hindlimb coordination.

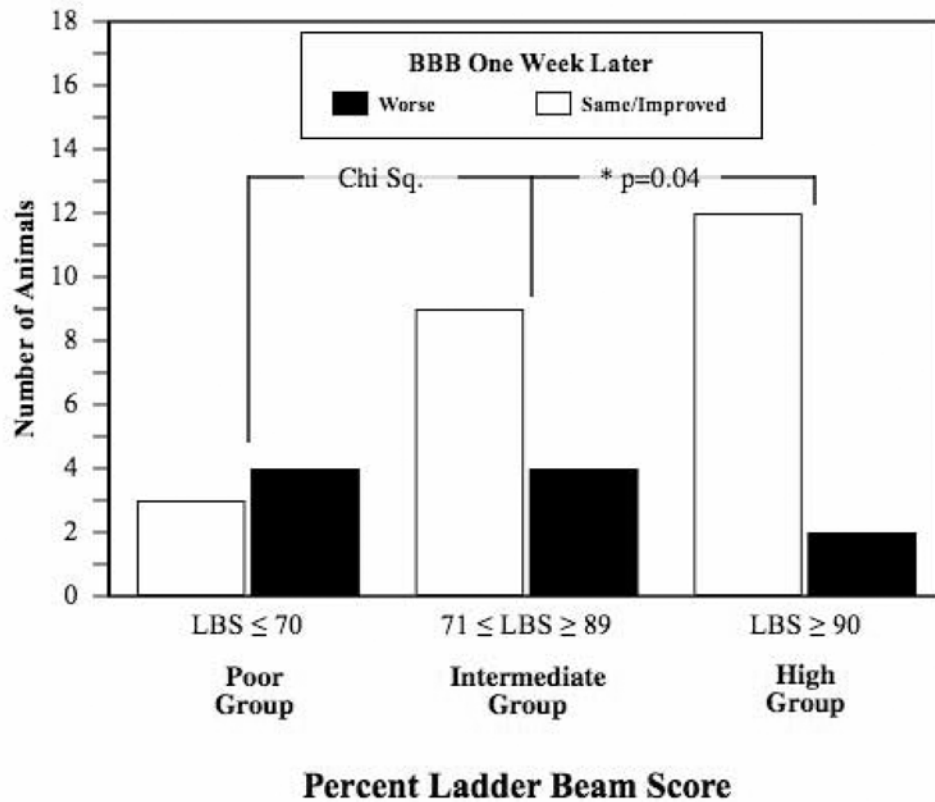


Figure 6. The percent Ladder Beam Score discriminates between animals that scored similarly on the BBB

Animals that scored between a 9 and 13 on the modified BBB at 14 days post-injury were also assessed on the ladder beam and then re-evaluated one week later on the BBB (n=34). Animals were grouped into poor, intermediate and high performers on the ladder beam task. Less than 50% of animals that received a percent LBS of 70 or less were the same or better on the BBB one week later (3 of 7). 69 percent of animals that received a percent LBS between 71 and 89 were the same or better on the BBB one week later (9 of 13), while 85% of animals that scored 90% or better on the LBS showed improvement on the BBB (12 of 14). A Chi square test for a trend across these three levels of LBS indicates a significance effect (Chi-Square = 4.06 (df 1), p=0.04).

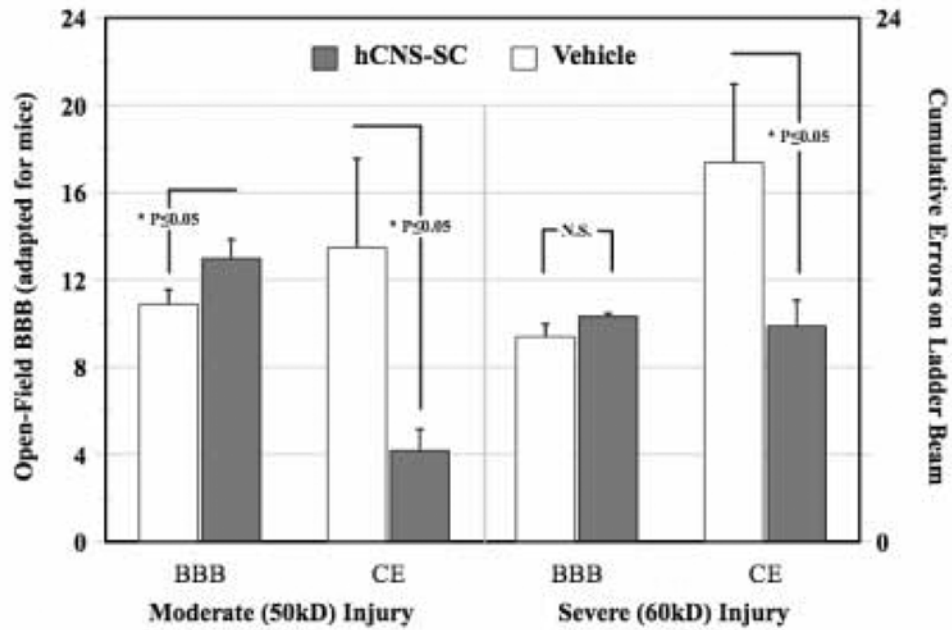


Figure 7. Differences between vehicle and stem cell treated animals detected by Cumulative Errors on the Ladder Beam

Animals received either a moderate (50kD) or more severe (60kD) SCI contusion injury followed by vehicle or human stem cell treatment 9 days post-injury. Only in the moderate injury paradigm was the BBB able to differentiate between groups. In both the moderate and severe injury paradigms, cumulative errors (CE) on the Ladder Beam task were able to differentiate between vehicle-treated versus hCNS-SC treated- groups 4-months post-grafting (*Student's T-test, $p \leq 0.05$, adapted from Cummings PNAS, 2005, used with permission).

Table 1**Operational definitions for the mouse modified BBB used for open-field assessment of locomotion in this project**

This scale is based on the BBB originally designed by Basso, Beattie and Bresnahan [1] and modified by Dergham and colleagues [7]. Since our observers are able to distinguish between trunk instability and trunk stability, we expanded Dergham's 17 point scale to 18 points. This modified mouse BBB scale has now been replaced by the Basso Mouse Scale (BMS) [3,8] and the modified scale should no longer be used.

Scale	Definitions for the BBB modified for use in mice
0	No observable hind limb movement
1	Slight movement of 1 or 2 joints of the hindlimb
2	Extensive movement of 1 joint of the hindlimb
3	Extensive movement of 2 joints of the hindlimb
4	Slight movement of a 3 joints of the hindlimb
5	Slight movement of 2 joints and extensive movement of a 3 rd
6	Extensive movement of 2 joints and slight movement of a 3 rd
7	Extensive movement of all 3 joints
8	Plantar placement of the paw with no weight support
9	Plantar placement of the paw with weight support in stance only
10	Occasional weight supporting stepping, no fore to hindlimb coordination
11	Frequent to consistent weight supported stepping, no fore to hindlimb coordination
12	Frequent to consistent weight supported stepping and occasional fore to hindlimb coordination
13	Frequent to consistent weight supported stepping and frequent fore to hindlimb coordination
14	Consistent plantar stepping, consistent coordination, paw rotation on initial contact and at lift off
15	Consistent plantar stepping, consistent coordination, parallel paw on initial contact and rotated at lift off
16	Consistent plantar stepping, consistent coordination, parallel paw on initial contact and at lift off, trunk instability and tail down
17	Consistent plantar stepping, consistent coordination, parallel paw on initial contact and at lift off, trunk instability, and consistent tail up
18	Consistent plantar stepping, consistent coordination, parallel paw on initial contact and at lift off, consistent trunk stability, and consistent tail up

Table 2**Operational definitions for positive events (green) and negative events (red) observed in mice during horizontal ladder beam crossing**

Each rung is analyzed, frame-by-frame, and is assigned a letter to denote one of six possible categories in a spreadsheet. When in doubt, a given rung is scored to the deficit: the animal is given a negative score. This increases the possibility of detecting of improvement over time. A data entry spreadsheet automatically calculates the ladder beam score (LBS %) as the number of positive events/total number of events and the cumulative number of errors across three trials (CE).

- (P) Plantar step.** The plantar surface of the hind paw makes initial contact with the rung, the paw does not need to fully grasp the rung. Toes are visible “above” (*i.e.* in the direction of forward movement) the rung throughout the majority of an active step cycle.
- (T) Toe step.** Only the digits make contact with the rung, no toes are observed “above” the rung until the last few frames of video corresponding to that step cycle (can allow for more frames if animal is moving very slowly).
- (S) Skip.** The hind paw skips over the rung during a normal, active step cycle. No part of the paw makes contact with the rung.
- (L) Slip.** Either the toes or plantar surface of the foot makes contact with the rung, but balance is lost and the foot drops below the plane of the rungs (*i.e.* from underneath vantage of the camera, the foot could be seen partially blocking view of the next rung).
- (M) Miss.** The hind paw passes between two rungs without touching. A step cycle was attempted but no contact of the rung being scored was made. The next rung was scored impartially from the low-point of the step cycle (*i.e.* a Miss between rungs 5 and 6 is attributed to rung 5 only, rung 6 is scored independently).
- (D) Drag.** The dorsal surface of the hind paw contacts the rung, or an incomplete step cycle occurs, or the hindlimb is in spasm while crossing the rung.