Published in final edited form as:

Gait Posture. 2014 February ; 39(2): 810–815. doi:10.1016/j.gaitpost.2013.11.006.

The intrarater reliability and agreement of compensatory stepping thresholds of healthy subjects

Jeremy R. Crenshaw, PhD and

Motion Analysis Laboratory, Division of Orthopedic Research, Mayo Clinic, Rochester, MN 55905 USA; Contact information: Mayo Clinic, 200 1st St. SW, Charlton North L-110K, Rochester, MN 55905, Telephone: 507 284 2262, Fax: 507 266 2227

Kenton R. Kaufman, PhD PE

Motion Analysis Laboratory, Division of Orthopedic Research, Mayo Clinic, Rochester, MN 55905 USA; Contact information: Mayo Clinic, 200 1st St. SW, Charlton North L-110K, Rochester, MN 55905, Telephone: 507 284 2262, Fax: 507 266 2227

Jeremy R. Crenshaw: crenshaw.jeremy@mayo.edu; Kenton R. Kaufman: kaufman.kenton@mayo.edu

Keywords

falls; clinical assessments; balance; stability; older adults

Introduction

In the United States in 2011, falls were the leading cause of unintentional non-fatal injuries, accounting for more than 9.2 million or 30.8% of all injuries [1]. When the population was stratified into 10 year age increments, falls were the leading cause of unintentional, non-fatal injuries for every age group except for those aged 15-24, for whom it ranked second. Falls are especially problematic for clinical populations with neuromuscular impairments, with notable yearly fall incidences observed for older adults (52% [2]), patients with Parkinson's disease (68% [3]), multiple sclerosis patients (64% [4]), and stroke survivors (43-70%, [5]). As the first step to preventative interventions for fall-related injuries, valid and reliable fall-risk assessments must be developed and implemented.

Fall-risk assessments that evaluate the response to external disturbances (e.g. a surface translation) have the potential to be valid because they specifically evaluate the neuromuscular processes necessary to recover from falls outside the laboratory. For example, trips and slips account for 59% of older adult falls in the community [2]. The successful recovery from trips [6; 7] and slips[8], as well as the recovery from surface translations [9-12], requires a coordinated response of a short reaction time, sufficient joint moments, and maintenance of the body center of mass (COM) within the base of support. Controlling the COM and maintaining dynamic stability can be accomplished with a feet-in-place response or a stepping response, depending on the situational constraints and disturbance direction and magnitude [13]. The ability to prevent a step, minimize the number of steps, or avoid a fall declines with age [14; 15], is worse for older adults fallers [16-19], and declines with Parkinson's disease [20; 21] and stroke [22-24]. Although such evidence supports *validity* of surface translation assessments, the *reliability* of such assessments is not known.

The purpose of this study was to evaluate the test-retest, intrarater reliability and agreement of compensatory stepping thresholds. Compensatory stepping thresholds are defined here as the minimum disturbance magnitude that consistently elicits a single or second compensatory step. As an initial step to establish the reliability of this assessment, a single

evaluator evaluated the response of young, healthy adults. We hypothesized that the evaluated assessments would show almost perfect test-retest reliability ($ICC > 0.80$) [25].

Methods

As per sample size recommendations for assessing reliability [26], a convenience sample of fifteen healthy adults under the age of 45 volunteered for this study (Table 1). Exclusion criteria included chronic or current pain, as well as self-reported medical conditions or previous injuries, trauma, or surgeries that negatively affect balance, mobility, or strength. Subjects did not use prostheses or assistive devices such as a cane or walker. Subjects had normal or corrected vision and were not knowingly pregnant. The study was approved by the Mayo Clinic Institutional Review Board, and all subjects provided written informed consent prior to participation.

Subjects visited the laboratory twice, with six to eight days between visits (6.9 ± 0.5 days). For each visit, subjects wore their own comfortable, well-cushioned athletic shoes. All testing occurred as subjects stood on a computer-controlled treadmill (ActiveStep®, Simbex, Lebanon, NH). Subjects were outfitted with a harness (Maine Anti-Gravity Systems, Inc., Portland, ME) attached to an overhead rail. The harness was adjusted so that, with full body weight support by the harness, the patient's knees could come close to, but not touch, the treadmill.

Each visit consisted of, in order, three progressive series of disturbances to determine anteroposterior single-stepping thresholds, anteroposterior multiple-stepping thresholds, and lateral stepping thresholds. During all progressions, 400 ms surface translations were delivered as the subject stood on the treadmill (Figure 1). The velocity profiles were triangular waveforms with peak velocities of 0.1 – 3.2 m/s, resulting in displacements of 2 – 64 cm and initial accelerations of 0.5 – 16.0 m/s².

Subjects initially assumed a standing upright posture before receiving a disturbance (Figure 1). During progressions to establish anteroposterior thresholds, subjects stood facing the front of the treadmill with feet placed side-by-side at a comfortable width (17.2 ± 4.4 cm). During the progression to establish lateral single-stepping thresholds, subjects stood facing the side of the treadmill with their feet placed touching together. Stance width was minimized for lateral disturbances as a safety measure to reduce the disturbance magnitude necessary to induce a step. Within each progression, the direction and the timing of the disturbances were pseudo-randomized to limit anticipatory improvements to the response. The disturbance direction (i.e. the direction of the fall) was randomized with the limitation that, at most, three disturbances of the same direction were consecutively delivered. Pseudo-randomization was achieved using a custom software macro (Excel 2010, Microsoft Corporation, Redmond, WA) that ordered subgroups of disturbances based on randomly generated numbers.

Disturbances were evenly assigned initial onset delays (i.e. the time after the investigator initiated the disturbance in software to when the belt moved) of 3 to 10 seconds.

Instructions were provided to constrain the stepping response. Before establishing anteroposterior single-stepping thresholds, subjects were instructed to “try not to step.” Failed responses were characterized by a step. Before establishing anteroposterior multiple-stepping thresholds, subjects were instructed to “try to take only one step”. Failed responses were characterized by more than one compensatory step (i.e. moving the non-stepping foot location or taking a subsequent step with the stepping foot). Before establishing the lateral single-stepping thresholds, subjects were instructed to “try not to step.” Failed responses were characterized by a step that extended the base of support in the direction of the

disturbance (i.e. the limb opposite to the falling direction crossed over or behind the stance foot or the limb on the same side as the falling direction moved laterally). For all disturbance types, responses were also identified as failures if the subject reported assistive support from the harness or the investigator observed unambiguous harness support. In the case of a failed response (e.g. Figure 1), the investigator told the subject the reason for the failure. All outcomes were evaluated by a single observer (JRC) with previous experience in evaluating the occurrence of compensatory steps. The evaluator was aware that ratings would be evaluated for reliability. In an attempt to maintain independence between ratings, the evaluator did not review the results of each evaluation until the end of the study.

Disturbance progressions were designed with the intent of minimizing variance due to aberrant responses and within-session learning. The first disturbance given for single-stepping thresholds had an absolute displacement of 2 cm, a peak speed of 10 cm/s, and an absolute initial acceleration of 50 cm/s². For multiple-stepping thresholds, the initial disturbance had an absolute displacement of 4 cm (peak speed = 20 cm/s; initial acceleration = 100 cm/s²). If a subject responded as instructed, the disturbance was increased by one “level” for the next disturbance in that direction. For single-stepping threshold progressions, levels were incremented by 50 cm/s² absolute initial accelerations (i.e. 2 cm absolute displacements and 10 cm/s peak speeds). For multiple-stepping threshold progressions, levels were incremented by 100 cm/s² initial accelerations (i.e. 4 cm displacements and 20 cm/s peak speeds). If the subject failed to respond as instructed, the subsequent disturbance in that direction was of the same magnitude. The disturbance that represented the threshold of interest elicited four consecutive failed responses. The disturbance progression continued until thresholds in each direction were established. If a threshold was established in one direction before the opposite direction, then the disturbance magnitude for the former direction was maintained at a level one less than the threshold. Lateral thresholds were identified relative to the mobilization side (i.e. preferred kicking limb) and stability side (i.e. preferred stance limb when kicking) [27].

Based on a simple inverted pendulum model, thresholds were expressed as the torque (τ) necessary to counter the torque caused by the disturbance (Figure 2):

$$\tau = |m \cdot \ddot{x} \cdot l|$$

where m is the subject mass, $\ddot{x}$ is the initial acceleration of the treadmill belt, and l is the estimated length of an inverted pendulum representing the standing subject ($l = 0.586 \cdot \text{height}$) [28]. This calculation characterizes the disturbance relative to the subject's mass and height.

The test-retest reliability of thresholds was assessed using custom software (MATLAB®, The MathWorks, Inc., Natick, MA) to calculate intraclass correlations (ICC (2,1)) [29]. Between-visit agreement was assessed using Bland-Altman plots and the limits of agreement (mean $\pm$ 1.96-standard deviation of Visit 2 – Visit 1) [30].

Results

Anterior single-stepping thresholds ($\tau = 273.0 \pm 82.3$ N·m) demonstrated almost perfect reliability (ICC (2,1) = 0.87, 95% C.I. = 0.66 – 0.95). The mean [limits of agreement] of the difference between visits was 10.2 [-75.1 to 95.4] N·m (Figure 3A). Fourteen of 15 differences (93%) were within the limits of agreement. *Posterior single-stepping thresholds* ($\tau = 235.5 \pm 98.0$ N·m) also demonstrated almost perfect reliability (ICC (2,1) = 0.97, 95% C.I. = 0.91 – 0.99). The mean [limits of agreement] of the difference between visits was 10.5

[-34.9 to 56] N·m (Figure 3B). All but two subjects maintained or improved their posterior threshold on the second visit (Table 1). All differences were within the limits of agreement.

Anterior multiple-stepping thresholds demonstrated almost perfect reliability (ICC (2,1) = 0.94, 95% C.I. = 0.70 – 0.99). Six subjects established anterior multiple-stepping thresholds on both visits ($\tau = 977.0 \pm 416.3$ N·m). Nine subjects did not establish a threshold, instead successfully recovering from the largest disturbance ($\ddot{x} = 16.0$ m/s², Table 1) in one step. The mean [limits of agreement] of the difference between visits was -62.7 [-349.7 to 224.2] N·m (Figure 3C). *Posterior multiple-stepping thresholds* also demonstrated almost perfect reliability (ICC (2,1) = 0.94, 95% C.I. = 0.83 – 0.98). Fourteen subjects established posterior multiple-stepping thresholds on both visits ($\tau = 701.9 \pm 237.5$ N·m). As a precaution against knee injury, the investigator ended the second visit of one subject (Table 1). This subject demonstrated a long step length and lower extremity orientation that an investigator determined to be an uncommon, yet potential risk for injury. The mean [limits of agreement] of the difference between visits was 21.9 [-142.6 to 186.5] N·m (Figure 3D). All but one difference was within the limits of agreement.

Mobility-side single-stepping thresholds ($\tau = 236.8 \pm 85.4$ N·m) demonstrated almost perfect reliability (ICC (2,1) = 0.95, 95% C.I. = 0.87 – 0.98). The mobility side for 13 of 15 subjects was their right side. The mean [limits of agreement] of the difference between visits was 11.0 [-37.3 to 59.3] N·m (Figure 3E). All but one subject maintained or improved their threshold on the second visit. Fourteen of 15 differences (93%) were within the limits of agreement. *Stability-side single-stepping thresholds* ($\tau = 225.7 \pm 77.7$ N·m) also demonstrated almost perfect reliability (ICC (2,1) = 0.97, 95% C.I. = 0.90 – 0.99). The mean [limits of agreement] of the difference between visits was 8.6 [-27.8 to 45.1] N·m (Figure 3F). All but one subject maintained or improved their threshold on the second visit. Fourteen of 15 differences (93%) were within the limits of agreement.

Discussion

To our knowledge, this is the first study to evaluate the test-retest, intrarater reliability and agreement of compensatory stepping thresholds. We hypothesized that the evaluated assessments would show almost perfect test-retest reliability (ICC > 0.80). This hypothesis was supported for all thresholds.

The demonstrated reliability confirms that this assessment is appropriate for research and clinical applications. Although reliability coefficients of 0.80 and higher are recommended for research-based group comparisons, higher reliability (ICC > 0.90) should be achieved for assessments that inform clinical decisions about an individual [25]. Based on this criterion, all thresholds but anterior single-stepping thresholds (ICC = 0.87) can be used confidently for patient assessments. Anterior multiple-stepping thresholds should also be used with caution, as the reliability was calculated from a small sample of subjects (n = 6) that established thresholds on both visits. The issue of recovery from the largest disturbance won't be as prevalent with patient populations. In unpublished data from an ongoing study of women over 65 years of age, no subjects progressed to the largest disturbance ($\ddot{x} = 8.4 \pm 2.5$; $\tau = 561.6 \pm 186.6$; n = 32). In the present study, the majority of discrepancies between visits were of one disturbance level (Table 1). Although the threshold levels below which high-risk individuals perform have not yet been identified, a potential within-subject error of one disturbance magnitude level should be considered when evaluating patient scores near such cutoff levels.

Reporting thresholds in units of torque, as opposed to units that describe the surface translation (e.g. displacement), gives biomechanical meaning to observed thresholds. The

simple model described here (Figure 2) assumes that an instantaneous torque is produced so that no excursion of the pendulum occurs. Such torque production is neither realistic nor required for successful responses. The single-stepping thresholds for anterior and posterior disturbances were approximately 2-3 and 5-10 times larger than the peak isometric plantarflexion and dorsiflexion torques for young adults, respectively [31]. The relative threshold values for plantarflexion ($4.0 \pm 0.7 \text{ N}\cdot\text{m}\cdot\text{kg}^{-1}$) and dorsiflexion ($3.4 \pm 1.0 \text{ N}\cdot\text{m}\cdot\text{kg}^{-1}$) were much higher than plantarflexion ($1 - 2 \text{ N}\cdot\text{m}\cdot\text{kg}^{-1}$) and dorsiflexion ($< 1 \text{ N}\cdot\text{m}\cdot\text{kg}^{-1}$) torques measured in response to instrumented surface translations [11]. Ongoing research is being conducted to correlate thresholds to isometric strength for older adult women. Previous research, however, suggests that thresholds and isometric strength will not be strongly correlated [32].

Of all thresholds, anterior single-stepping thresholds demonstrated the poorest reliability ($\text{ICC} (2,1) = 0.87$). This observation may be a result of greater within-subject variance due to multiple recovery strategies. Postural responses can be characterized by a continuum of muscle synergies that span from an ankle strategy to a hip strategy, increasing inter-trial variability [33]. The variance in available strategies may not be as large for stepping responses and other recovery directions. Protocol modifications should be considered to improve the reliability for these assessments. First, reliability may be improved by adding additional observers, video recordings that are immediately available for replay, and objective measures of stepping. Employing an instrumented harness to objectively evaluate harness support may also improve the reliability for cases in which harness support is in question. A method for determining a cutoff level of harness support (20% body weight) has been described in the literature [34]. This method was evaluated for forward lean-releases, but a criterion has not been evaluated for surface translations. The number of failed trials needed to establish a threshold may also be altered from the four-trial criterion held in this study. We do not recommend increasing this number, as this will increase the number of trials necessary to establish thresholds. For anteroposterior single-stepping thresholds, anteroposterior multiple-stepping thresholds, and lateral single-stepping thresholds, subjects experienced 26.3 ± 5.1 , 40.2 ± 5.6 , and 22.8 ± 5.4 trials, respectively. For many patient populations, a large number of trials will induce a confounding variable of fatigue. Reducing the number of failed trials that defines a threshold may reduce reliability and not accurately measure patient capabilities.

Single and multiple compensatory-stepping thresholds join a list of objective functional fall-risk assessments with acceptable reliabilities for making clinical decisions ($\text{ICC} > 0.90$). These tests include the functional reach test ($\text{ICC} = 0.92$), performance-based home assessment ($\text{ICC} = 0.92$), lateral reach ($\text{ICC} = 0.99$), timed up and go ($\text{ICC} = 0.56 - 0.99$), Morse fall scale ($\text{ICC} = 0.96$), Tinetti balance subscale ($\text{ICC} = 0.98$), and Berg balance scale ($\text{ICC} = 0.98$) [35]. Combining multiple assessments may improve the validity of the measure without sacrificing reliability. For example, the Balance Evaluation Systems Test (BESTest, interrater $\text{ICC} = 0.91$) consists of 36 items, grouped into systems of biomechanical constraints, stability limits/verticality, anticipatory postural adjustments, postural responses, sensory orientation, and stability in gait [36]. The clinical utility of these measures is a product of the assessment's reliability, validity, and feasibility.

Compensatory stepping threshold assessments lack a degree of ecological validity. Here, we define ecological validity as the extent to which the assessments measure the ability to recover from common fall causes (i.e. a trip or slip [2]) in the free-living environment. Ecological validity is a compromise between mimicking real-life circumstances while maintaining scientific control, upholding patient safety, and sustaining patient participation. Such validity is a product of the assessment setting, stimuli, and response [37]. Trips and slips likely occur while ambulating overground. An initial standing position on a treadmill,

however, allowed for mediolateral disturbances and eliminated the potentially confounding variance in gait kinematics and disturbance timing relative to the phase of gait. The surface translation stimulus is similar to trips and slips in that an external stimulus is applied to the distal lower extremities, creating postural instability by displacing the body COM with respect to the base of support and necessitating an active response to avoid a fall. Therefore, the proposed assessments have greater ecological validity than previously established assessments that challenge stability with self-initiated disturbances (e.g. functional reach). Furthermore, the reliability of compensatory stepping thresholds appears to be greater than that of the perturbation-based Sensory Organization Tests (ICC = 0.09-0.81 [38]).

The ecological validity of the response is limited by the constrained instructions of “try not to step” and “try to take only one step”. Stepping constraints alter the preferred stepping response [39]. Although single-step responses may not be necessary to recover from trips and slips, the first step mechanics are important, discriminant aspects to trip and slip recovery [6; 40]. The first-step kinematics that are important to fall recovery, such as step length, step width, step time, and reduced trunk rotation are likely the same variables important to the recovery from a surface translation in a single step. If the number of steps were not constrained, the likely variance in the incidence and number of steps would limit the assessment reliability and clinical utility.

While non-stepping responses may not mimic the stepping response attempted in 45% of falls outside the laboratory [41], feet-in-place responses have a degree of similarity to stepping responses. Both strategies necessitate a coordinated, direction-specific, and timely response of the lower-extremity musculature that can be augmented by upper extremity motion to maintain the body COM within the base of support [42-47]. The advantage of feet-in-place assessments, which are accompanied by smaller disturbance magnitudes than multiple-stepping assessments (Figure 3 and Table 1), is improved patient adherence and safety. In an ongoing study of 88 older adult women, 3 patients did not complete the anteroposterior single-stepping threshold assessment due to anxiety. Six patients did not complete the multiple-stepping threshold assessment due to anxiety (4) or discomfort of the knee or back (2).

Additional research is needed to establish stepping thresholds as a clinical measure. First, the intra- and interrater reliability should be established for each population of interest. Although the results of this study suggest that the measures have excellent reliability, both the within- and between-patient variability may be affected by old age or pathology. Second, the population-specific sensitivity and specificity of predicting future falls needs to be established for stepping thresholds. It is plausible that specific thresholds may determine risk from certain types of falls (i.e. *anterior* thresholds predict *trip-induced falls*, and *posterior* thresholds predict *slip-induced falls*). To our knowledge, assessments specific to the type of fall have not been evaluated, and our ongoing studies are designed to evaluate this prospect.

Acknowledgments

External Funding Source: Funding for JRC provided by NIH T32 HD0744. The information or content and conclusions do not necessarily represent the official position of, nor should any official endorsement be inferred by the National Institutes of Health.

References

1. Centers for Disease Control and Prevention. 10 Leading Causes of Nonfatal Injury, United States--2011, Office of Statistics and Programming, National Center for Injury Prevention and Control, CDC using WISQARS™.

2. Berg WP, Alessio HM, Mills EM, Tong C. Circumstances and consequences of falls in independent community-dwelling older adults. *Age Ageing*. 1997; 26:261–268. [PubMed: 9271288]
3. Wood B, Bilclough J, Bowron A, Walker R. Incidence and prediction of falls in Parkinson's disease: a prospective multidisciplinary study. *Journal of Neurology, Neurosurgery & Psychiatry*. 2002; 72:721–725.
4. Peterson EW, Cho CC, von Koch L, Finlayson ML. Injurious Falls Among Middle Aged and Older Adults With Multiple Sclerosis. *Arch Phys Med Rehabil*. 2008; 89:1031–1037. [PubMed: 18503796]
5. Weerdesteyn V, De Niet M, Van Duijnhoven HJR, Geurts ACH. Falls in individuals with stroke. *J Rehabil Res Dev*. 2008; 45:1195–1214. [PubMed: 19235120]
6. Pavol MJ, Owings TM, Foley KT, Grabiner MD. Mechanisms leading to a fall from an induced trip in healthy older adults. *J Gerontol A Biol Sci Med Sci*. 2001; 56:M428–437. [PubMed: 11445602]
7. Pijnappels M, Bobbert MF, van Dieen JH. Push-off reactions in recovery after tripping discriminate young subjects, older non-fallers and older fallers. *Gait Posture*. 2005; 21:388–394. [PubMed: 15886128]
8. Redfern MS, Cham R, Gielo-Periczak K, Grönqvist R, Hirvonen M, Lanshammar H, Marpet M, Pai CYC, Powers C. Biomechanics of slips. *Ergonomics*. 2001; 44:1138–1166. [PubMed: 11794762]
9. Crenshaw JR, Rosenblatt NJ, Hurt CP, Grabiner MD. The discriminant capabilities of stability measures, trunk kinematics, and step kinematics in classifying successful and failed compensatory stepping responses by young adults. *J Biomech*. 2012; 45:129–133. [PubMed: 22018682]
10. Dimitrova D, Horak FB, Nutt JG. Postural Muscle Responses to Multidirectional Translations in Patients with Parkinson's Disease. *Journal of Neurophysiology*. 2004; 91:489–501. [PubMed: 12944541]
11. Hall CD, Woollacott MH, Jensen JL. Age-related changes in rate and magnitude of ankle torque development: implications for balance control. *J Gerontol A Biol Sci Med Sci*. 1999; 54:M507–513. [PubMed: 10568533]
12. Jensen JL, Brown LA, Woollacott MH. Compensatory stepping: the biomechanics of a preferred response among older adults. *Exp Aging Res*. 2001; 27:361–376. [PubMed: 11681198]
13. McIlroy W, Maki B. Adaptive changes to compensatory stepping responses. *Gait Posture*. 1995; 3:43–50.
14. Hall CD, Jensen JL. Age-related differences in lower extremity power after support surface perturbations. *J Am Geriatr Soc*. 2002; 50:1782–1788. [PubMed: 12410895]
15. Mille ML, Rogers MW, Martinez K, Hedman LD, Johnson ME, Lord SR, Fitzpatrick RC. Thresholds for inducing protective stepping responses to external perturbations of human standing. *Journal of Neurophysiology*. 2003; 90:666–674. [PubMed: 12711707]
16. Hilliard MJ, Martinez KM, Janssen I, Edwards B, Mille ML, Zhang Y, Rogers MW. Lateral balance factors predict future falls in community-living older adults. *Arch Phys Med Rehabil*. 2008; 89:1708–1713. [PubMed: 18760155]
17. Mansfield A, Maki BE. Are age-related impairments in change-in-support balance reactions dependent on the method of balance perturbation? *J Biomech*. 2009; 42:1023–1031. [PubMed: 19362311]
18. Pai YC, Rogers MW, Patton J, Cain TD, Hanke TA. Static versus dynamic predictions of protective stepping following waist-pull perturbations in young and older adults. *J Biomech*. 1998; 31:1111–1118. [PubMed: 9882043]
19. Rogers MW, Hedman LD, Johnson ME, Cain TD, Hanke TA. Lateral stability during forward-induced stepping for dynamic balance recovery in young and older adults. *J Gerontol A Biol Sci Med Sci*. 2001; 56:M589–594. [PubMed: 11524454]
20. Jacobs JV, Dimitrova DM, Nutt JG, Horak FB. Can stooped posture explain multidirectional postural instability in patients with Parkinson's disease? *Experimental Brain Research*. 2005; 166:78–88. [PubMed: 16096779]
21. McVey MA, Amundsen S, Barnds A, Lyons KE, Pahwa R, Mahnken JD, Luchies CW. The effect of moderate Parkinson's disease on compensatory backwards stepping. *Gait and Posture*. 2013 in press. 10.1016/j.gaitpost.2013.03.028

22. Lakhani B, Mansfield A, Inness EL, McIlroy WE. Compensatory stepping responses in individuals with stroke: A pilot study. *Physiotherapy Theory and Practice*. 2011; 27:299–309. [PubMed: 20958215]
23. Marigold DS, Eng JJ. Altered timing of postural reflexes contributes to falling in persons with chronic stroke. *Experimental Brain Research*. 2006; 171:459–468. [PubMed: 16418855]
24. Marigold DS, Eng JJ, Timothy Inglis J. Modulation of ankle muscle postural reflexes in stroke: influence of weight-bearing load. *Clinical Neurophysiology*. 2004; 115:2789–2797. [PubMed: 15546787]
25. Nunnally, JC.; Bernstein, IH. *Psychometric theory*. McGraw; New York: 1994. 1991
26. Fleiss, JL. *Design and analysis of clinical experiments*. Wiley-Interscience; 2011.
27. Sadeghi H, Allard P, Prince F, Labelle H. Symmetry and limb dominance in able-bodied gait: a review. *Gait and Posture*. 2000; 12:34–80. [PubMed: 10996295]
28. Clauser CE, McConville JT, Young JW. Weight, volume, and center of mass of segments of the human body, DTIC Document. 1969
29. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979; 86:420–428. [PubMed: 18839484]
30. Martin Bland J, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *The Lancet*. 1986; 327:307–310.
31. Moraux A, Canal A, Ollivier G, Ledoux I, Doppler V, Payan C, Hogrel JY. Ankle dorsi-and plantar-flexion torques measured by dynamometry in healthy subjects from 5 to 80 years. *BMC Musculoskelet Disord*. 2013; 14:104. [PubMed: 23522186]
32. Mille ML, Creath RA, Prettyman MG, Johnson Hilliard M, Martinez KM, Mackinnon CD, Rogers MW. Posture and locomotion coupling: a target for rehabilitation interventions in persons with Parkinson's disease. *Parkinsons Dis*. 2012; 2012:754186. [PubMed: 22295253]
33. Torres-Oviedo G, Ting LH. Muscle synergies characterizing human postural responses. *Journal of Neurophysiology*. 2007; 98:2144–2156. [PubMed: 17652413]
34. Cyr MA, Smeesters C. Maximum allowable force on a safety harness cable to discriminate a successful from a failed balance recovery. *J Biomech*. 2009; 42:1566–1569. [PubMed: 19446297]
35. Scott V, Votova K, Scanlan A, Close J. Multifactorial and functional mobility assessment tools for fall risk among older adults in community, home-support, long-term and acute care settings. *Age Ageing*. 2007; 36:130–139. [PubMed: 17293604]
36. Horak FB, Wrisley DM, Frank J. The balance evaluation systems test (BESTest) to differentiate balance deficits. *Phys Ther*. 2009; 89:484–498. [PubMed: 19329772]
37. Schmuckler MA. What is ecological validity? A dimensional analysis. *Infancy*. 2001; 2:419–436.
38. Ford-Smith CD, Wyman JF, Elswick R, Fernandez T, Newton RA. Test-retest reliability of the sensory organization test in noninstitutionalized older adults. *Arch Phys Med Rehabil*. 1995; 76:77–81. [PubMed: 7811180]
39. McIlroy WE, Maki BE. Task constraints on foot movement and the incidence of compensatory stepping following perturbation of upright stance. *Brain Res*. 1993; 616:30–38. [PubMed: 8358619]
40. Troy KL, Donovan SJ, Marone JR, Bareither ML, Grabiner MD. Modifiable performance domain risk-factors associated with slip-related falls. *Gait Posture*. 2008; 28:461–465. [PubMed: 18396048]
41. Holliday, PJ.; Fernie, G.; Gryfe, CI.; Griggs, GT. *Slips, Stumbles, and Falls: Pedestrian Footwear and Surfaces (ASTM STP 1103)*. Philadelphia: American Society for Testing and Materials; 1990. Video recording of spontaneous falls of the elderly; p. 7-16.
42. Allum J, Carpenter M, Honegger F, Adkin A, Bloem B. Age-dependent variations in the directional sensitivity of balance corrections and compensatory arm movements in man. *The Journal of physiology*. 2002; 542:643–663. [PubMed: 12122159]
43. Henry SM, Fung J, Horak FB. Control of stance during lateral and anterior/posterior surface translations. *Rehabilitation Engineering, IEEE Transactions on*. 1998; 6:32–42.

44. Marigold DS, Bethune AJ, Patla AE. Role of the unperturbed limb and arms in the reactive recovery response to an unexpected slip during locomotion. *Journal of Neurophysiology*. 2003; 89:1727–1737. [PubMed: 12611998]
45. Pijnappels M, Kingma I, Wezenberg D, Reurink G, van Dieën JH. Armed against falls: the contribution of arm movements to balance recovery after tripping. *Experimental Brain Research*. 2010; 201:689–699. [PubMed: 19949781]
46. Schillings A, Van Wezel B, Mulder T, Duysens J. Muscular responses and movement strategies during stumbling over obstacles. *Journal of Neurophysiology*. 2000; 83:2093–2102. [PubMed: 10758119]
47. Tang PF, Woollacott MH. Inefficient postural responses to unexpected slips during walking in older adults. *The Journals of Gerontology Series A: Biological Sciences and Medical Sciences*. 1998; 53:M471–M480.

**Anterior/
Posterior
Single
Stepping
Thresholds** **Anterior/
Posterior
Multiple
Stepping
Thresholds** **Lateral
Single
Stepping
Thresholds** **Figure 1.**

Figure 2.

Figure 3.

Table 1

Individual subject data

Subject	Sex	Age (years)	Mass (kg)	Pendulum Height (m)	Stepping Threshold Disturbance Initial Acceleration (m/s^2)									
					Anterior Single		Posterior Single		Anterior Multiple		Posterior Multiple		Mobility-Side Single	
					Visit 1	Visit 2	Visit 1	Visit 2	Visit 1	Visit 2	Visit 1	Visit 2	Visit 1	Visit 2
1	F	28	66.0	1.01	3.5	3.0	1.0	0.5	11.0	12.0	7.0	7.0	2.0	2.0
2	M	31	71.6	1.05	4.0	4.5	4.0	4.0	MAX	16.0	13.0	15.0	3.0	3.0
3	F	22	55.2	0.95	4.0	4.0	3.0	3.5	14.0	8.0	10.0	7.0	2.5	2.0
4	M	23	78.2	1.08	4.5	5.0	4.0	4.0	MAX	MAX	12.0	14.0	4.0	4.0
5	M	28	102.0	1.09	3.0	4.0	3.5	3.5	16.0	16.0	9.0	10.0	3.5	2.5
6	M	23	74.6	1.05	4.0	4.5	3.5	3.5	MAX	MAX	15.0	NA	2.5	3.5
7	F	25	44.6	0.90	4.0	3.5	3.0	3.5	MAX	MAX	11.0	11.0	3.0	3.5
8	F	33	56.6	0.96	3.5	4.0	2.5	3.0	MAX	MAX	9.0	9.0	3.0	3.0
9	F	40	58.1	0.98	3.0	3.5	3.0	2.5	13.0	12.0	9.0	9.0	2.5	2.5
10	F	24	66.7	1.04	3.0	3.0	3.5	4.0	16.0	14.0	11.0	11.0	3.5	4.0
11	F	42	53.5	0.97	4.5	4.5	3.5	3.5	16.0	MAX	12.0	11.0	4.0	4.0
12	F	44	67.3	1.00	4.0	4.5	3.0	3.5	14.0	15.0	9.0	9.0	3.5	3.0
13	M	23	82.0	1.05	4.5	4.0	4.5	5.0	MAX	MAX	10.0	11.0	4.5	4.5
14	F	22	68.4	1.00	3.5	3.5	3.5	3.5	MAX	16.0	10.0	10.0	3.5	4.0
15	F	27	67.0	1.05	5.5	4.5	3.5	4.0	MAX	MAX	10.0	10.0	4.5	4.0

Note: Pendulum height is 58.6% of body-height [22]. MAX indicates that the subject recovered from the largest disturbance ($\hat{x} = 16 m/s^2$) with one step. NS indicates that the protocol was ended before a threshold was established. Thresholds were calculated from body mass (m), estimated pendulum length (l), and initial disturbance acceleration ($\hat{x}$): $t = |m \cdot \hat{x} / l|$.