2010

The ABLE Scale: The Development and Psychometric Properties of a New Outcome Measure for the Spinal Cord Injury Population

Elizabeth Ardolino
Seton Hall University

The ABLE Scale: The Development And Psychometric Properties Of A New Outcome Measure For The Spinal Cord Injury Population

By

Elizabeth Ardolino

Dissertation Committee:

Dr. Genevieve Pinto Zipp, Chair

Dr. MaryAnn Clark

Dr. Karen Hutchinson

Approved by the Dissertation Committee:

Genevieve Pinto Zipp Date: 3-20-10

MaryAnn Clark Date: 3-20-10

Karen J Hutchinson Date: 3-19-2010

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Health Sciences

Seton Hall University

2010

Acknowledgements

As I learned rather early in this journey, the development of a new outcome measure is a huge undertaking. I am fortunate to have had the help of many generous and talented people. To all of those listed below, I owe my sincerest gratitude.

To my dissertation committee for their undying patience and support. To Dr. Genevieve Pinto Zipp, who showed me where to begin, and how to end. To Dr. MaryAnn Clark, for her fire-cracker way of keeping our discussions on track. To Dr. Karen Hutchinson, who never stopped asking the hard questions, and who challenged me to make the ABLE scale great.

To the members of the NRN Balance Committee for reading 14 versions of the ABLE scale before the first patient was ever tested on it, and for providing the clinical and financial support necessary to accomplish this study.

To the many data collectors and physical therapists at Magee Rehabilitation Hospital, Frazier Rehabilitation, The Shepherd Center, and Kessler Rehabilitation, who generously donated their time and talents to collect data on over 100 subjects in 5 months.

To all of my colleagues at Magee, especially: Elizabeth Watson, Carol Owens, Melania Dieterick, Alice Kennedy and Mary Schmidt, who not only helped me with data collection, but who helped maintain my sanity on a daily basis.

To Dr. Mike Linacre, Dr. Craig Velozo, Dr. Elaine Walker and Dr. Sergio Romero, for all of their statistical assistance with the Rasch analysis.

To Andrea Behrman and Susie Harkema, whose passion for promoting recovery after SCI inspired me.

To my parents, Frank and Thomasine Roberts, for their endless words of encouragement and financial support.

To my wonderful husband, Art, who taught himself how to program the Rasch analysis software, who helped me with all things computer-related, and who never doubted that I would finish this project.

Dedication

To Meredith: for being not only my sister, but my first best friend.

To Art: for being not only my husband, but my soul mate.

Table of Contents

List of Tables

List of Figures

Abstract

The ABLE Scale: The Development and Psychometric Properties of a New Balance Outcome Measure for the Spinal Cord Injury Population.

Elizabeth Ardolino, PT, MS

Seton Hall University

May 2010

Chair: Genevieve Pinto Zipp

Objectives: To develop and examine the initial psychometric properties of a new balance outcome measure for the spinal cord injury population, using a Rasch analysis.

Design: This exploratory research study utilized a methodological research design to test the initial psychometric properties of a new balance outcome measure for the SCI population, the ABLE scale. The properties tested were targeting, item difficulty, and person separation reliability.

Setting: Four outpatient and inpatient rehabilitation hospitals.

Participants: A total of 104 individuals with non-progressive spinal cord injuries.

Main Outcome Measures: The Activity-based Balance Level Evaluation (the ABLE scale).

Results: Initial analysis of the ABLE scale using Rasch analysis revealed floor and ceiling effects and multiple item redundancies. Applying pivot anchoring to the analysis resulted in the improved targeting range of the scale, although there are still several items with similar difficulty levels, indicating a redundancy of these items. The person separation of the ABLE scale was calculated to be 7.67, with a person separation reliability of .98.

Conclusion: This was the first step in the development and testing of a new balance outcome measure for the spinal cord injury population. The Rasch analysis provided a method for identifying changes that need to be made to the ABLE scale. Future studies are needed to further test the psychometric properties of this new scale.

Chapter I

INTRODUCTION

Background of the Problem

A spinal cord injury is a sudden, catastrophic, life-changing event. There are an estimated 12,000 new cases of spinal cord injury (SCI) each year in the United States (*Spinal cord injury: Facts & figures at a glance.* 2008) . Presently, it is estimated that there are greater than 1.2 million individuals living with a SCI in the U.S. (*One degree of separation: Paralysis and spinal cord injury in the United States* 2009).

The SCI most often results in some degree of sensation and/or motor loss below the level of the lesion. The severity of the SCI ranges from a complete injury, in which an individual has no sensation or movement at or below the level of the lesion, to an incomplete injury, in which an individual has some motor and/or sensation below the level of the lesion (Somers, 2000). This loss of voluntary movement and sensation often results in an impairment in balance, which impacts the injured individual's ability to perform functional activities, such as rolling in bed, transferring in and out of bed, standing, walking, or propelling a manual wheelchair. These impairments also result in difficulty with performing activities of daily living, such as dressing, bathing, eating, and toileting (Somers, 2000).

In the SCI population, an individual's balance, in both sitting and standing, largely affects their mobility and participation in functional activities. Both mobility and independence have been shown to impact an individual's quality of life. Studies have shown that individuals with greater participation in their community, such as having a job, going to school, or being able to drive, were more satisfied with their lives than individuals who were not able to play a role in their community (Charlifue & Gerhart, 2004; Dijkers, 2005; Franceschini, Di Clemente, Rampello, Nora, & Spizzichino, 2003): Thus, individuals with SCI who have better balance and mobility will have greater participation in their community, resulting in a higher quality of life.

Individuals with SCI have an increased incidence of injury (Krause, 2004). Krause (2004) found that 19% of the 1328 subjects with

SCI studied sustained at least one injury within a 12 month period. Brotherton et al (2007) found that seventy-five percent of the 119 individuals with incomplete SCI who were surveyed reported at least one fall in the past year. These individuals cited a loss of balance as one of the major contributors to their fall.

The measurement of balance is an essential component of the evaluation process, as it guides the clinician in establishing a prognosis and a plan of care, assists in the assessment of the effectiveness of an intervention, aides in the design of an appropriate wheelchair seating system, and guides the clinician in determining an appropriate assistive device for ambulation. While it is obvious that assessing balance in an individual with SCI is important, there are few objective means by which to do this. In the literature, researchers studying balance of individuals with SCI utilize force plates and EMG surface electrodes to measure changes in center of pressure and muscle activation (Grigorenko et al., 2004; Y. J. M. Janssen-Potten, Seelen, Drukker, Huson, & Drost, 2001; Y. J. M. Janssen-Potten, Seelen, Drukker, Spaans, & Drost, 2002; Y. Janssen-Potten, Seelen, Drukker, & Reulen, 2000; Kamper et al., 1999; Middleton, Sinclair, Smith, & Davis, 1999; Shirado, Kawase, Minami, & Strax, 2004). While these measures provide precise and objective data, they cannot be utilized in the typical physical therapy clinical setting. Force plates are costly to purchase and install, and require a dedicated space in a well-controlled environment. In addition, personnel must be highly-trained in order to reliably collect and analyze data collected from both force plates and EMG (Monsell, Furman, Herdman, Konrad, & Shepard, 1997).

As force plates and EMG are not available in the typical physical therapy clinic, physical therapists have utilized clinical outcome measures to assess balance in the SCI population. Three outcome measures often used by clinicians are Berg Balance Scale (BBS) (K. Berg, Wood-Dauphinee, Williams, & Gayton, 1989), the Tinetti Performance-Oriented Mobility Assessment (POMA)(Tinetti, 1986), and the Modified Functional Reach Test (MFRT)(Lynch, Leahy, & Barker, 1998).

The BBS (K. Berg et al., 1989) is useful for assessing balance in the portion of the SCI population who have recovered the ability to stand. The test consists of 14 items which test both static and dynamic functional standing balance. Many of the BBS items are timed to allow for more objective scoring. Furthermore, the test requires little equipment, most of which can be found in the typical physical therapy clinic or a client's home. The test also does not require lengthy training before it be used to assess a client.

Only two studies have examined the reliability and validity of the BBS in the SCI population (Wirz et al., 2009; Lemay & Nadeau, 2009). Wirz et al., (2009) found excellent interrater reliability, and established the concurrent validity of the BBS with several other outcome measures. However, this study only tested these properties in the chronic SCI population, and only the concurrent validity was established for this outcome measure. Lemay and Nadeau (2009) established the concurrent validity of the BBS with the Walking Index for Spinal Cord Injury (WISCI II), the Spinal Cord Injury Functional Ambulation Index (SCI-FAI), the 10-meter walk test, and the Timed-Up and Go test (TUG). However, these authors found a ceiling effect with the BBS, and recommended that it be used in conjunction with other outcome measures. Furthermore, the BBS has only been used in 2 other published research studies in the SCI population (Behrman & Harkema, 2000; Dobkin, Apple, Barbeau, & et al., 2006). In their 2000 case series of four subjects with incomplete SCI, Behrman and Harkema reported the use of the BBS in one of their four subjects. These authors expressed concerns that the BBS had not yet been validated for the SCI population. In a 2006 multi-center randomized clinical trial of locomotor training in the incomplete SCI population, Dobkin et al., (2006) utilized the BBS as a secondary outcome measure. The results of the BBS were not reported in this -publication (Dobkin et al., 2006).

Thus, the greatest weakness of using the BBS in the SCI population is the paucity of its known psychometric properties for this specific population. As the psychometric properties of an outcome measure are dependent upon the population being tested, the scarceness of these established properties in the SCI population is a considerable drawback to its use in the clinical and research settings.

Further weaknesses of the BBS for the SCI population include the testing items themselves. The BBS consists of only 1 sitting balance item. Thus, the test has an inherent floor effect for a large percentage of the SCI population, who are non-ambulatory, primary wheelchair users. On the other end of the spectrum of recovery, there is a percentage of individuals with SCI who regain the ability to ambulate without an assistive device. The test has a ceiling effect for these patients, as they have regained balance in the standing domain, but still present with high level balance deficits during ambulation. As the BBS does not have a gait component, this test is unable to capture these deficits in this population.

As the BBS was initially developed for the elderly population, several of the test items assume that the individual being tested has adequate upper extremity range of motion and function. In the SCI population, where over 50% of the injuries are in the cervical spinal cord, many individuals have impaired upper extremities which often results in decreased shoulder range of motion and strength, and impaired hand function, including the inability to grasp and pick up objects (*Spinal cord injury: Facts & figures at a glance.*2008). Thus, several of the BBS items, which require the ability to grasp an object or elevate the upper extremity, are not applicable in the SCI population, without modifying the scoring or administration of the item.

Another balance outcome measure, the Performance-oriented Mobility Assessment (POMA) (Tinetti, 1986) is utilized by clinicians treating patients with SCI for several reasons. First, the POMA balance component has fewer items than the BBS and therefore can often be administered in less time than the BBS. Second, the POMA has one item that tests reactions to an external perturbation (a nudge applied in standing), an indicator of reactive or anticipatory balance response, which has been identified as important contributor to balance responses. Thirdly, the POMA has a gait subscale, which can be useful in identifying balance deficits and/or changes in balance in patients who are ambulatory with and without an assistive device. Fourth, the test requires little equipment, most of which can be found in the typical physical therapy clinic or a client's home. The test does also not require a large amount of training in order to use with a patient, as demonstrated by Cipriany-Dacko et al. (1997).

As with the BBS, the POMA has several weaknesses when being administered with the SCI population. The psychometric properties of the POMA in the SCI population have not been reported in any published studies. In fact, there are no published studies on the SCI population that have utilized the POMA as an outcome measure. As the psychometric properties of an outcome measure are dependent upon the population being tested, the lack of these established properties in the SCI population is a huge drawback to its use in the clinical and research settings.

Further weaknesses of the POMA lie also in the testing items themselves. As in the BBS, the POMA only consists of one sitting balance item. Therefore, the POMA has a floor effect for individuals with SCI who are non-ambulatory, primary wheelchair users. While the POMA does have a gait subscale, the majority of the items on this scale assess the quality of the gait, as opposed to the subject's abilities to perform functional tasks during gait.

Also, as the scoring of the items on the POMA is very vague (allowing for no performance of the item, abnormal performance, or normal performance), the POMA may lack the sensitivity to detect small, yet perhaps functional, changes in the ability to perform each item.

The Modified Functional Reach Test (MFRT) (Lynch et al., 1998) is another outcome measure often utilized by physical therapists treating individuals with SCI. This test differs from the BBS and POMA in that it is the only outcome measure which has been specifically designed for use in the SCI population. The reliability of this outcome measure has been established in the motor complete SCI population (Lynch, Leahy, & Barker, 1998). The strength of the MFRT is that it can be used in patients who are non-ambulatory and are unable to stand. Also, this test is easy and quick to administer, and requires minimal equipment.

Unfortunately, very little is known regarding the psychometric properties of the MFRT in the SCI, or any, population. There are only two published studies on the properties of this test (Lynch, Leahy, & Barker, 1998, Adegoke et al., 2002). Both of these studies were conducted on very homogenous populations, consisting of young males with motor complete SCI who were able to elevate their upper extremity to 90 degrees of shoulder flexion (Lynch, Leahy, & Barker, 1998, Adegoke et al., 2002). Therefore, we do not know how well these results would generalize to other individuals with SCI, especially those with impaired upper extremity use, who are unable to achieve 90 degrees of shoulder flexion. Furthermore, the inter-rater reliability, the validity, and the minimal detectable change need to be established in this population.

Another weakness of the MFRT is that it consists of only one test item, which is measured three times. While this does give the clinician some idea of how the patient's sitting balance is progressing, it is not a complete assessment of the patient's functional abilities in sitting. Thus, the MFRT may be limited in its sensitivity to assess changes in balance in this population.

In summary, the means by which researchers have measured balance in the SCI population, mainly with the use of force plates and EMG data, are unavailable in the typical physical therapy clinic. There are currently no outcome measures that have been developed and validated to specifically assess balance in the SCI population. The clinical outcome measures that are often utilized, including the Berg Balance Scale, the Tinetti Performance-Oriented Mobility Assessment, and the Modified Functional Reach Test, have not be validated for use in the SCI population. Also, each of these tests address slightly different aspects of balance assessment and performance, and each have inherent weaknesses. Therefore, there is a need for a new balance outcome measure specific to the SCI population. Thus, the first purpose of this study is to develop an all-inclusive, valid clinical instrument, the Activity-based Balance Level Evaluation (ABLE scale) that assesses balance across the full spectrum of recovery in the SCI population. The second purpose of this study is to determine the initial psychometric properties of the ABLE scale in the SCI population.

Research Questions:

Question 1: Does the ABLE scale have the appropriate range of item difficulty to capture balance abilities across the spectrum of recovery in the SCI population?

Question 2: Does the ABLE scale exhibit the psychometric properties of targeting, item difficulty, and person separation reliability?

Chapter II

REVIEW OF THE LITERATURE

The review of the literature is divided into three main sections. The first section is a description of balance, including a depiction of how balance is often assessed in the SCI population. The second section provides the reader with background information on current clinical outcome measures used to assess balance. This section emphasizes the development and psychometric properties of three clinical outcome measures. The third section describes the methodology behind the development of a new clinical outcome measure.

Description of Balance

The ability to maintain one's balance is a key element in successfully performing functional activities. Balance is often defined as the ability to maintain one's center of gravity (COG) over one's base of support (BOS) (Allison & Fuller, 2001). Balance is mediated by the interaction of the visual, vestibular, and somatosensory systems

(Forssberg & Nashner, 1982). The visual system provides information on one's orientation in the environment, as well as detects one's motion through that environment. The vestibular system provides information regarding the position and movement of one's head in relation to gravity. This system provides the central nervous system (CNS) with the information necessary to differentiate self-motion from movement in the environment. The CNS uses information from somatosensory receptors located in joints, ligaments, muscles, and skin to provide information on the motions of the feet in relation to the contact forces of the support surface (Nashner, Black, & Wall, 1982).

The interaction of these three systems creates the ability of the body to maintain stable positions, automatically respond to voluntary postural changes, and appropriately react to external perturbations (K. Berg et al., 1989). These abilities are possible through automatic and anticipatory postural responses, and volitional postural movements (Allison & Fuller, 2001). Automatic postural responses are long loop reflexes that work to maintain a stable position and react to external perturbations. There are 4 commonly identified automatic postural responses. The ankle strategy is defined as the use of the ankles and feet to control postural sway. The hip strategy is the use of the pelvis and trunk to control postural sway. The suspensatory strategy is the control of balance by squatting to lower the COG over the BOS. Finally, the stepping strategy employs movement of the feet to establish a new BOS (Allison & Fuller, 2001). While the automatic postural responses are used to react to a disturbance after it occurs, anticipatory postural responses are used in advance to respond to a predicted disturbance. These anticipatory responses allow an individual to maintain his balance while catching a baseball or lifting a heavy suitcase. Volitional postural movements allow one to control balance during self-initiated movements, and can range from weight shifting during ambulation, to reaching up to a high shelf, to hitting a volley in a game of tennis. An inability to control any of these aspects of balance may limit one's functional activities as well as result in an increased risk for falls (Allison & Fuller, 2001). Furthermore, separating out sensory integration issues from maladaptive motor responses due to a disordered movement system is critical in identifying the etiology of a fall.

The measurement of balance gives some indication of an individual's risk of falling (Cattaneo et al., 2002; Chern, Yang, & Wu, 2006; Garland, Willems, Ivanova, & Miller, 2003; Gavin-Dreschnack et al., 2005; J. Smith, Forster, & Young, 2006). Individuals with neurologic disorders, such as cerebrovascular accidents (CVA), traumatic brain injury, spinal cord injury (SCI), or multiple sclerosis (MS), often are at risk of falling. Research on risk of falls in these populations has shown that individuals with acute CVA have a fall risk range of 10.5-46% (Smith et al., 2006), SCI wheelchair users have a 37.9% chance of falling and a 3-17% chance of a serious wheelchair accident resulting in injury (Gavin-Dreschnack et al., 2005), and individuals with MS have a 54% risk of falls (Cattaneo et al., 2002). While the fall is rarely fatal, it frequently results in injuries requiring costly and lengthy hospitalizations, as well as decreases in participation in functional activities (Cattaneo et al., 2002; Gavin-Dreschnack et al., 2005; Smith et al., 2006). While measuring balance is a good assessment of fall risk, balance has been used as both an impairment measure and as a measure of physical function and mobility in intervention studies (Bastille & Body, 2004; Langhammer & Stanghelle, 2003). Balance has also been measured to assess the effectiveness of an intervention, to measure functional recovery after neurologic disorders, and as a source of variance in ambulation activity (Eng, Chu, Dawson, Kim, & Hepburn, 2002; Eng et al., 2003; Michael, Allen, & Macko, 2005; Rochester et al., 2004).

The measurement of balance is an important consideration in the spinal cord injury population. According to the National Spinal

Cord Injury Statistical Center, there are an estimated 12,000 spinal cord injury (SCI) cases each year (*Spinal cord injury: Facts & figures at a glance*. 2008). Assessment of balance in this population is essential in order to determine a person's ability to safely perform functional activities, such as bed mobility, transfers, and ambulation. An individual's balance will also play a role in the design of an appropriate wheelchair seating system. Ultimately, measuring balance in this population may help to predict the ability to stand and ambulate, as well as predict the risk of falling.

In the SCI population, an individual's sitting and standing balance largely affects their mobility and level of independence. Both mobility and independence have been shown to impact an individual's participation and quality of life. Studies have shown that individuals with greater community integration, such as having a job, going to school, or being able to drive, were more satisfied with their lives than individuals who were not able to play a role in their community (Charlifue & Gerhart, 2004; Dijkers, 1999; Franceschini, Di Clemente, Rampello, Nora, & Spizzichino, 2003). This implies that balance may affect the participation, and thus, the quality of life of individuals with SCI.

Only a few studies have measured balance in the SCI population (Grigorenko et al., 2004; Y. J. M. Janssen-Potten, Seelen, Drukker, Huson, & Drost, 2001; Y. J. M. Janssen-Potten, Seelen, Drukker, Spaans, & Drost, 2002; Y. Janssen-Potten, Seelen, Drukker, & Reulen, 2000; Kamper et al., 1999; Middleton, Sinclair, Smith, & Davis, 1999; Shirado, Kiwase, Minami, & Strax, 2004). All but one of these studies assessed balance in only a sitting position. Measuring sitting balance is important in the SCI population, as the ability to maintain both static and dynamic sitting balance affects an individual's independence with activities of daily living, transfers, and the propulsion of wheelchairs (Grigorenko et al., 2004; Y. J. M. Janssen-Potten, Seelen, Drukker, Huson, & Drost, 2001; Y. J. M. Janssen-Potten, Seelen, Drukker, Spaans, & Drost, 2002; Y. Janssen-Potten, Seelen, Drukker, & Reulen, 2000; Kamper et al., 1999; Shirado, Kiwase, Minami, & Strax, 2004). Janssen-Potten et al (2000, 2001, 2002) reported three separate studies examining the effects of different wheelchair configurations on the balance of individuals with complete thoracic or lumbar SCI. These authors assessed balance during a choice reaction time task by measuring changes in center of pressure (COP) using force plates, as well as, changes in muscle activity using EMG surface electrodes (Y. J. M. Janssen-Potten et al., 2001; Y. J. M. Janssen-Potten et al., 2002; Y.

Janssen-Potten et al., 2000). The researchers found that while the type of footrest and inclination of the seat did not significantly affect the balance of individuals with thoracic injuries, the COP displacement was increased when these individuals were seated in more customized chair configurations. This indicated that individuals with complete injuries relied on this customization to allow for improved use of accessory muscles in order to maintain their balance (Y. J. M. Janssen-Potten et al., 2001; Y. J. M. Janssen-Potten et al., 2002; Y. Janssen-Potten et al., 2000). Force plates were also used in the examination of long sitting in patients with complete thoracic SCI (Shirado et al., 2004). Shirado et al (2004) found that when subjects with complete SCI raised their arms over their thighs while in long sitting, they experienced a posterior shift in COP. This contrasted with the anterior COP shift noted in able-bodied subjects. The authors reasoned that this posterior COP shift was a way to counteract the anterior instability that resulted from the arm elevation in the SCI subjects. Kamper et al (1999) utilized force plates in the assessment of postural stability during external perturbations that simulated driving conditions, in individuals who are manual wheelchair users. The authors found that the majority of the SCI subjects became unstable during all levels of the external perturbations, and used compensatory methods to reduce COP displacement (Kamper et al., 1999). Grigorenko et al (2004) measured the effects of kayak training on sitting balance in individuals with SCI by using force plates. The training consisted of paddling an open-air kayak with a seat especially designed for individuals with SCI, two to three times per week, for 8 weeks. To test sitting balance, ground reaction and COP forces were recorded for each subject while sitting quietly with eyes open and knees slightly bent, in a specially-designed chair. Based upon the COP measurements, these authors found that kayak training had only a minimal effect on the sitting balance of these subjects, although 75% of the subjects reported subjective feelings of improved sitting balance in their wheelchairs (Grigorenko et al., 2004). Therefore, one might question whether the methodology for assessing balance in this study was appropriate. Perhaps the authors would have seen more significant changes had they measured functional balance, as opposed to just measurements of COP, given that during the kayak training, subjects were moving their trunk and upper extremities, and not maintaining a stationary position.

Middleton, Sinclair, Smith, and Davis (1999) were the only authors to study standing balance. These investigators also used a force platform to measure the effect of medially-linked knee-ankle-foot orthoses (KAFOs) on postural stability and sway during quiet standing and functional tasks with and without one-handed support, in individuals with SCI (Middleton, Sinclair, Smith, & Davis, 1999). According to the authors, a low mean amplitude of sway indicated high postural stability, and therefore better balance control. The authors found that wearing KAFOs that were linked together significantly improved postural control, as defined by changes in sway as measured by force plates, when compared to wearing unlinked KAFOs (Middleton, Sinclair, Smith, & Davis, 1999).

While this previous work provides a better understanding of sitting balance in the SCI population, the results are limited in their generalizability. All of these studies used subjects with motor complete thoracic and/or lumbar SCI, thus ignoring a large percentage of the SCI population. In 2007, 52.4% of all SCI injuries reported to the University of Alabama database from the SCI Model Systems Centers were in the cervical spinal cord, and only 23.0% were complete injuries in the thoracic and lumbar regions (*Spinal cord injury: Facts & figures at a glance.*2008). Therefore, these studies have only provided us information on 23% of the SCI population. It is important to capture data on balance in individuals with SCI at all 3 regions of the spinal cord, as well as in both complete and incomplete injuries. There may be a large amount of variability in the balance of an individual with a complete cervical injury, versus an incomplete cervical injury, as well as between individuals with complete vs. incomplete paraplegia. However, at this point in time, no one has measured these differences across varying neurological and severity levels, nor has anyone tracked changes in balance function in these individuals over time.

Measuring Balance

The studies by Janssen-Potten et al. (2000,2001, 2002), Grigorenko et al. (2004), Kamper et al. (1999), and Shirado et al. (2002) do provide good examples of the type of data that can be collected by using force plates to assess the balance of individuals with SCI. In computerized dynamic platform posturography (CDDP) a mechanized force platform is used to assess the visual and somatosensory inputs to balance in standing, and is often applied in clinical research studies (Monsell et al., 1997)(Monsell, Furman, Herdman, Konrad, & Shepard, 1997). Computerized dynamic platform posturography can be utilized during the Sensory Organization Test (SOT). The SOT employs moveable forceplates and a moveable visual surround system to alter the somatosensory and visual environments. During this testing, the platforms can be moved horizontally and the visual stimuli can be manipulated by removing vision (i.e. having the subject close his/her eyes), or through the moveable visual surround. The SOT measures a subject's amount of sway while standing quietly during six different conditions (eyes open with steady platform; eyes closed on steady platform; moving visual field on steady platform; eyes open on sway-referenced moveable platform; eyes closed on moveable platform; and moving visual field on movable platform) (Allison & Fuller, 2001; Monsell et al., 1997). While CDDP and the SOT have been found to be reliable and valid, the use of force plates in the typical physical therapy clinic is not always feasible (Allison & Fuller, 2001; Monsell et al., 1997). Force plates are costly to purchase and install, and require a dedicated space in a well-controlled environment. In addition, personnel must have a high level of expertise to reliably collect and analyze the force plate data. Furthermore, it is often unclear how the data produced by the force plates correlates to functional activities, as was demonstrated by Grigorenko et al (2004). Therefore to address these issues, in 1986, Shumway-Cook and Horak developed the Clinical Test for Sensory Interaction on Balance (CTSIB). This test uses the same six conditions as the SOT, however the forceplates are replaced with a foam pad, and a stopwatch and visual observation are used to collect measurements. One drawback in using the CTSIB, as well as the SOT, in individuals with SCI is that the test must be performed in standing. Thus, individuals with SCI who are unable to stand cannot be assessed with these tests.

Three commonly used clinical outcome measures are the Berg Balance Scale (BBS) (K. Berg et al., 1989), the Tinetti Performance-Oriented Mobility Assessment (POMA) (Tinetti, 1986), and the Modified Functional Reach Test (MFRT) (Lynch et al., 1998). These three outcome measures are often utilized in physical therapy clinics with patients with varying diagnoses. However, when using assessment tools, one must ascertain the psychometric properties of these tools. Psychometric properties include reliability (intra-rater and inter-rater), validity (concurrent, convergent, divergent, construct, predictive), effect size, and floor and ceiling effects. These properties provide insight into the correlation of the tools with other outcome measures previously shown to be valid for a specific population (concurrent validity), the correlation of the tools with other tests believed to measure the same phenomenon (convergent validity), the disassociation of the tools with other tests believe to measure different phenomenon (divergent validity), the generalizability of these tools (construct validity), and whether the scales predict scores on other measures (predictive validity). Effect size provides insight into the clinical significance of the scales by showing how large or small a change can be expected from the scale. Floor (minimal possible score achieved) and ceiling (maximal possible score achieved) effects help determine the utility of the tool in a specific population. The purpose of this section of the literature review is to examine the development, psychometric properties, and current usage of these three outcome measures, and address the rationale for their use in the SCI population. The results of this review will support the need to develop a new tool to assess balance in these individuals.

Development of the Berg Balance Scale

The Berg Balance Scale (BBS) was developed as a clinical tool for measuring balance in the elderly population (K. Berg et al., 1989). The BBS was designed to meet three goals: serve as a means of predicting falls in the elderly population, evaluate the effectiveness of a balance intervention, and aid therapists in establishing a plan of care (K. Berg et al., 1989). In their initial study, the authors described the three phases of scale development, as well as the initial testing of reliability and criterion validity (K. Berg et al., 1989).

The purpose of the first phase was to develop a large group of items considered for inclusion in this instrument (K. Berg et al., 1989).

Ten health care professionals, including physical and occupational therapists, physicians, and nurses, were asked to describe the movements and actions that they felt fully assessed a person's balance. In addition, twelve elderly patients answered open-ended questions about the activities that made them unsteady. Based on these answers, the investigators designed 38 items that were functional movements, the majority of which contained a time component to ensure objectivity of measurement (K. Berg et al., 1989).

The goal of the second phase was to decrease the number of items (K. Berg et al., 1989). A test was conducted to assure clarity and appropriateness, which eliminated 5 items. 11 professionals were then asked to determine the value of each item. 14 patients performed the 33-item test, and rated their perception of steadiness on each item. Items were dropped from the test if the professionals deemed them unimportant, or if the item did not discriminate balancing ability among patients. Eleven items were eliminated from the test at this phase (K. Berg et al., 1989).

The third phase focused on item elimination and verifying the content of the scale (K. Berg et al., 1989). At this stage, 12 patients were videotaped while performing each of the 22 items on the test.

The participants were asked to rate their sense of steadiness for each item, and these perceptions were compared to the actual scores. Content validity was established based on the consistency of these comparisons (K. Berg et al., 1989).

The authors assessed interrater and intrarater reliability. Interrater reliability, which tests the agreement between raters, was assessed with 5 testers and 14 patients. Intrarater reliability assesses the agreement between ratings taken by the same rater. The authors measured reliability using the intraclass correlation coefficient (ICC), which estimates the average correlation between all possible pairs of ratings. A perfect correlation has an ICC of 1.00. The interrater reliability ICC for the total scores was 0.98 (K. Berg et al., 1989). The intrarater reliability ICC was 0.99 (K. Berg et al., 1989). Internal consistency was calculated through a correlation matrix, in which each item was compared to all other items, and the total correlations ranged from 0.72 to 0.90 (K. Berg et al., 1989). None of the correlations suggested item redundancy.

Berg et al. (1989) established criterion validity by correlating the balance scores with the observed global ratings (i.e. subjective views of balance) given by treating therapists. The investigators found a significant correlation ($p<.01$). Thus, Berg and colleagues (1989) stated that the new balance scale was reliable and valid in the elderly population. As the psychometric properties of an outcome measure are specific to the population in which they are used, the authors recommended further testing of the validity in other populations.

Psychometric Properties of the Berg Balance Scale

Validity of the BBS in the elderly and CVA populations.

Several studies have examined the concurrent and predictive validity of the BBS in different populations (K. O. Berg, Maki, Williams, Holliday, & Wood-Dauphinee, 1992; K. O. Berg, Wood-Dauphinee, Williams, & Maki, 1992; Bogle Thorbahn & Newton, 1996; Harada et al., 1995; Juneja, Czyrny, & Linn, 1998; Liston & Brouwer, 1996; Mao, Hsueh, Tang, Sheu, & Hsieh, 2002; Stevenson & Garland, 1996; Teasell, McRae, Foley, & Bhardwaj, 2002; Wee, Wong, & Palepu, 2003). Berg et al (1992b) published the first study to assess the validity of the scale in the elderly and stroke populations. Concurrent validity in the elderly population was tested through correlations of the balance scores with the global ratings of the subjects' caregivers, and the self-perceptions of the individual subjects. These moderate correlations were statistically significant. In this same study, the BBS scores were also found to be able to distinguish between the need for different mobility aids. Individuals with higher BBS scores required less reliance on assistive devices. Further moderate, yet statistically significant, correlations between the scale scores and the performance of elderly individuals in laboratory measures of sway on a moving platform fitted with force plates, also supported the concurrent validity of the BBS. The ability of the scale to predict the risk of falls was also validated in this study. The number of falls that 93 elderly subjects experienced in 12 months was correlated with the scores on the BBS (K. O. Berg, Wood-Dauphinee, Williams, & Maki, 1992). The authors also tested the use of the scale in patients with stroke. The BBS was found to correlate highly with the Barthel Index (a measurement of functional independence with self-care and mobility post stroke) (Mahoney & Barthel, 1965) and the Fugel-Meyer Scale (a measurement of recovery of motor performance after stroke) (Fugel-Meyer, Jaasko, Leyman, Olsson, & Steglind, 1975) in 60 patients status-post stroke, who were followed over a 12-week period (K. O. Berg, Maki, Williams, Holliday, & Wood-Dauphinee, 1992). Correlations with the Barthel Index in these patients ranged from .80 at initial evaluation to .94 at discharge. Correlations with the Fugel-Meyer Scale ranged from .77 at initial evaluation to .82 at discharge. The authors cautioned that while these results established the validity of the BBS for the elderly and stroke populations, more research was needed by other investigators to confirm their results (K. O. Berg, Wood-Dauphinee et al., 1992b).

Berg and colleagues (1992a) established the concurrent validity of the BBS through correlations with the mobility section of the Barthel Index (r=.67), the balance subscale of Tinetti's Performance-oriented Mobility Assessment (r=.91), the Timed-Up and Go Test (r= -.76), the need for assistive devices (effect size >1), and laboratory measures of sway and response to perturbations (r= -.55) in 31 elderly subjects. The BBS was moderately to strongly correlated with all of the above measures. The BBS was also found to be the only test to statistically distinguish between subjects who walked with a walker, a cane, or no assistive device, which further supports earlier findings. The interpretations of the results of this study are limited, however, by the small sample size of 31 subjects (K. O. Berg et al., 1992a).

These studies validated the BBS' ability to assess two essential features of standing balance: the ability to maintain static stance, and the ability to maintain standing balance in response to external perturbations (K. O. Berg et al., 1992; K. O. Berg, Wood-Dauphinee et al., 1992). Stevenson and Garland (1996) examined the BBS' ability to assess a third feature of standing balance: the ability to make anticipatory postural adjustments to voluntary movements. The researchers examined the center of pressure excursion during self-initiated rapid arm flexion in 24 subjects with chronic stroke. These investigators found that scores on the BBS correlated highly (r= .81) with measurements of center of pressure, meaning that subjects with higher BBS scores exhibited increased ability to successfully accommodate to a perturbation. The authors argued that these results reflect the BBS' ability to test this third aspect of standing balance (Stevenson & Garland, 1996).

The reliability and validity of the BBS in the stroke population was assessed through two other studies (Liston & Brouwer, 1996; Mao et al., 2002). Liston and Brouwer (1996) reported excellent test-retest reliability (ICC=0.98) in a sample of 20 subjects with chronic stroke. Mao et al (2002) found an interrater ICC of 0.95 for the BBS in a sample of 112 patients followed for 180 days after stroke onset. These investigators also showed concurrent validity through high correlations of the BBS with the Fugel-Meyer test and the Postural Assessment Scale for Stroke Patients (Benaim, Perennou, Villy, Rousseaux, & Pelissier, 1999). Convergent validity was established through high correlations with the

Barthel Index, and predictive validity was confirmed through the walking subscale of the Motor Assessment Scale (Carr, Shepherd, Nordholm, & Lunne, 1985) (Mao et al., 2002).

The predictive validity of the BBS.

One important property of the BBS is its ability to predict falls in the elderly population, which was assessed by Bogle Thorbahn and Newton (1996). A multiple-regression analysis was used to determine how each factor reported on an activity index questionnaire related to the subject's score on the BBS. The authors found the BBS to have a high specificity (96%), but a low sensitivity (53%), meaning that the BBS was only able to identify 53% of the individuals who actually had a fall. Thus, the BBS was not accurate in identifying those subjects at risk for falls (Bogle Thorbahn & Newton, 1996). Interestingly, the researchers found that the subjects who fell more often were those with scores closer to the cutoff score of 45/56. The authors hypothesized that those subjects who scored in the lowest range of the test had developed compensatory strategies to minimize their number of falls (Bogle Thorbahn & Newton, 1996). This suggests that individuals scoring closer to the cutoff score of 45/56 may need more intensive balance training and further education on compensatory strategies for preventing falls.

These findings are disparate to the results of Harada et al (1995), who found a specificity level of 78% and a sensitivity level of 84% in a sample of 53 elderly subjects. The divergence in results may be the consequence of different methodologies. A receiver operating characteristic (ROC) curve helps determine how well a test discriminates between the presence and absence of a condition. Harada et al (1995) utilized a ROC curve to determine an optimum cutoff score of 48/56. It was at this point on the ROC curve that the authors found the best tradeoff between specificity and sensitivity (Harada et al., 1995).

A more recent study examined the ability of the BBS to predict multiple falls in 210 community-dwelling elderly individuals (Muir, Berg, Chesworth, & Speechley, 2008). The researchers determined an optimal cutoff score of 54/56, based upon a ROC analysis. These authors found that the previously-determined cutoff score of 45/56 was inadequate for predicting future fallers. The authors recommended that the BBS not be used as a dichotomous scale with a cutoff of 45/56 in this population.

While the previous two investigations examined the BBS's ability to predict falls in the elderly population, a 2002 study tested the predictive validity of the BBS in the acute stroke population (Teasell et al., 2002)(Teasell et al., 2002). Teasell et al (2002) found that in a sample of 238 patients with stroke in an acute rehabilitation facility, admission BBS scores were significantly lower for patients who fell during their course of rehabilitation (19.0/56 for fallers, vs. 30.7/56 for non-fallers). Based upon these findings, the BBS was recommended as a tool for rehabilitation professionals to identify patients who were at risk for falls (Teasell et al., 2002).

A recent study looked at several balance measures across the various domains of the International Classification of Functioning, Disability, and Health (ICF) model, and determined how accurately they could identify individuals with stroke who had a history of multiple falls (Beninato, Portney, & Sullivan, 2009). Using a ROC analysis, the researchers found that the BBS was only moderately able to identify multiple fallers (AUC=0.76). These authors also found a ceiling effect in this population of 27 individuals with chronic stroke.

The validity of the BBS in predicting outcomes for patients in acute rehabilitation was assessed in three studies? (Juneja et al., 1998; J. Y. Wee et al., 2003; J. Y. M. Wee et al., 1999). Juneja et al (1998) used a sample of 45 patients with a variety of diagnoses who were in an inpatient acute rehabilitation facility. The authors discovered that the BBS scores upon admission to acute rehabilitation had a moderate to high correlation with the Functional Independence Measure (FIM) (Juneja et al., 1998). The FIM has been shown to be a valid predictor of outcomes, and thus a high correlation with this measure indicates good predictive validity of the BBS (Juneja et al., 1998). The authors suggested that decreased balance scores do not limit the patient's potential for functional gains, but may be used as a moderator variable in determining prognosis and a plan of care (Juneja et al., 1998). Wee et al (2003, 1999) found that the admission BBS score in the acute stroke population was moderately correlated with length of stay (LOS) ($r = -.53$). They also used the BBS as a predictor of discharge destination, and found that patients with an admission BBS score above 20 were more likely to be discharged home ($p<.001$) (J. Y. Wee et al., 2003). The authors did not state what the discharge BBS score was of the patients who were discharged to home.

The ability of the BBS to detect changes in balance.

An important aspect of any outcome measure is the ability to accurately detect change over time (English & Hillier, 2006). In a study of 60 patients with acute stroke, stratified across three general functional levels of severity, the BBS was shown to be as responsive to change as the Barthel Index (Wood-Dauphinee, Berg, Bravo, & Williams, 1997). In this study, the BBS demonstrated a larger effect size than the Barthel Index or the Fugel-Meyer scale and also suggested that the BBS was capable of discriminating between patients of different levels of severity. Concomitantly, English et al (2006) also found the BBS to have a large effect size, suggesting that it was very sensitive to detecting changes in balance. A negligible floor effect and a minimal ceiling effect were seen in this study of 61 subjects with acute stroke. Although these authors did not stratify their patients according to disease severity, they suggested that the scale was appropriate for patients along a spectrum of balance abilities (English & Hillier, 2006).

Several statistical analyses are required to determine the effect size of a measure, making it impractical for clinicians to use effect sizes to make decisions about individual patients. Thus, it is helpful to know the minimal detectable change (MDC) of an instrument. The MDC is "the criterion amount of change that must occur in order for a clinician to conclude that genuine change has occurred (Stevenson, 2001)." Stevenson (2001) determined the MDC of the BBS in an acute stroke population to be 6 points. However, the author warned that the methodology used to determine the MDC might have overestimated this value. In this study, the patient's primary physical therapist first administered the test, allowing the patient to attempt each task 3 times, and scoring the patient's best attempt. Within 24 hours of this first rating, a second physical therapist re-assessed the patient with the BBS. The patient then engaged in physical therapy for one to two weeks, and was re-assessed by their primary physical therapist, who administered the BBS again, and also assigned a general rating of the degree of change in the patient's balance. Had a within-rater format been used instead of a between-rater format, the MDC might have been lower than 6 points. A more recent study by Liaw et al (2008) determined the smallest real difference (SRD) of the BBS in the chronic stroke population. In a study of 52 individuals with chronic stroke, the SRD was calculated to be 6.68 affirming a change of 7 points is necessary on the BBS for the clinician to conclude that the patient has improved his balance (Liaw et al., 2008). Two other studies established the smallest detectable difference (SDD) for the BBS in the Parkinson's and elderly populations (M. Conradsson, Lundin-Olsson, Lindelof, & et al., 2007; Lim, van Wegen, de Goede, & et al., 2005). Lim et al (2005) determined the SDD of the BBS to be 3 points in patients with PD.

Contrary to this, Conradsson et al (2007) recently established an SDD of 8 points in elderly adults living in residential care facilities. The dramatic differences seen between these three studies are most likely due to the differences in the population. Stevenson et al (2001) used 48 acute CVA patients with an average age of 73.5, and an average initial BBS score of 43.0. Lim et al (2005) studied 26 subjects with PD, with an average age of 62.5 years, an average Mini-Mental State Exam (MMSE) of 24, and an average initial BBS score of 53.8. In the Conradsson et al study (2007), 45 elderly adults who required assistance for activities of daily living, with an average age of 82.3 years, an average initial BBS score of 30.1, and an average MMSE score of 17.5 were studied. Furthermore, the latter two studies used an intrarater design, whereas Stevenson et al (2001) used an interrater design. Obviously one's diagnosis, age, cognition, and initial balance abilities play a role in determining the minimal detectable change on the BBS. Therefore, it is recommended that the MDC be established in each specific neurological population for which it is applied.

The validity of the BBS for individuals with multiple sclerosis.

The apparent utility of the BBS has lead many clinicians to use it with a variety of patient populations. While the majority of research on the psychometric properties of the BBS has been conducted on the elderly and stroke populations, several recent articles reported on the reliability and validity in other neurological populations. Paltaama et al (2005) discovered that the BBS had high intrarater and interrater reliability in a sample of 28 ambulatory patients with multiple sclerosis (MS). However, the researchers did find that there was a large ceiling effect for this population, as many of the patients scored a 56/56. Therefore they recommended that the BBS be used with other balance measures (Paltaama et al., 2005). Conversely, in a study of 51 patients with MS with varying ambulatory abilities, Cattaneo et al (2006) found a very low ceiling effect, as many patients were non-ambulatory. These authors established concurrent validity of the BBS in this population through moderate correlations with the Dynamic Gait Index (DGI), an 8-item scale that measures dynamic balance during functional mobility. The researchers also found that in the MS population, as in the well-elderly, when discriminating between fallers and non-fallers, the specificity of the BBS was high, but the sensitivity was low. This is in agreement with the results reported by Bogle Thorbahn and Newton (1996). Cattaneo et al (2006) reasoned that this low sensitivity may be caused by other impairments present in patients with MS that are not tested by the BBS. The disparity observed between the findings of these two articles most likely resulted from the differences in the population. The subjects in Paltaama et al.'s (2005) study had a mean disease duration of 5.8 years, and a mean BBS score range of 50 to 54.5. On the contrary, Cattaneo et al's (2006) subjects had a longer length of time since injury, with an average of 15.6 years, and a lower mean BBS score (47.5.) Therefore, given the shorter disease duration, and the higher mean BBS score, it is not surprising that Paltaama et al (2005) found a ceiling effect whereas Cattaneo et al (2006) did not.

Nilsagård et al (2009) examined the ability of the BBS to predict falls in 76 adults with MS. The researchers found that at a cutoff of ≤55/56, the BBS had a sensitivity of .96 and a specificity of .14. Thus, the BBS was able to correctly identify fallers in this population. However, the authors did find a ceiling effect in 13 of the subjects tested (Nilsagård et al., 2009).

The validity of the BBS for individuals with Parkinson's disease.

Three articles published in 2005 addressed the use of the BBS in patients with Parkinson's disease (PD) (Brusse, Zimdars, Zalewski, & Steffen, 2005; Franchignoni, Matignoni, Ferriero, & Pasetti, 2005; Qutubuddin et al., 2005). In 70 ambulatory patients with PD, Franchignoni et al (2005), found moderate correlations between the BBS, the Fear of Falling Measure (a 19-item self-administered questionnaire on concern of falling while performing activities of daily living)(Velozo & Peterson, 2001), and the Postural Changes Scale (a 3-item, 5-point ordinal scale that tests rolling, rising from the floor, and rising from a bed), indicating construct validity. The BBS also had an alpha coefficient of Cronbach of 0.95, indicating good internal consistency when used in this population (Franchignoni et al., 2005). Furthermore, Brusse et al (2005) determined the internal consistency on the items in the BBS was 0.88 in a population of 25 subjects with PD. These two reports of high internal consistency suggest that the items on the BBS all measure the same construct of balance. Brusse et al (2005) also established concurrent validity for this sample through good correlations of the BBS and the TUG (r= -.78), and the BBS and the forward functional reach test (r= .50). Qutubuddin et al (2005) tested criterion validity in 38 males with PD and found that lower BBS scores correlated with lower patient functioning on the Unified Parkinson's Disease Rating Scale. Higher BBS scores also correlated with lower Hoehn and Yahr Scale staging, which indicates fewer signs of the disease. The findings of these three studies indicate that the BBS is a valid tool for measuring balance in patients with PD (Brusse et al., 2005; Franchignoni et al., 2005; Qutubuddin et al., 2005).

Furthermore, in a more recent study, Steffen et al. (2008) determined the test-retest reliability of 37 subjects with PD , and found an ICC= .94. The authors also calculated the MDC of the BBS in this population, and found an MDC of 5 points. It should be noted that the patients tested in this study were very high-functioning, and had a mean score on the BBS of 50/56 (Steffen & Seney, 2008).

The psychometric properties of the BBS in the SCI population.

To date, very little literature has examined the psychometric properties of the BBS in the SCI population. Datta et al. (2009) used a principle component analysis to examine how each item on the BBS contributes to the variability of the entire scale. Data were analyzed from 97 patients with incomplete SCI (ASIA C and D) participating in locomotor training across five of the centers of the NeuroRecovery Network. The NeuroRecovery Network is a network of seven outpatient physical therapy clinics which provide standardized locomotor training to individuals with spinal cord injury to promote their neurological recovery (Datta et al., 2009). The patients were placed into one of three subgroups, referred to as Phases by these authors. Phase I patients were unable to stand or walk and were primary wheelchair users. Phase II patients were able to stand for limited amounts of time with or without assistive devices, but still relied on a wheelchair for their primary means of mobility. Phase III patients were able to ambulate with or without assistive devices (Datta et al., 2009). The investigators determined the first principal component, which consisted of all BBS items except item #3 (sitting balance), accounted for 48% of the total variability in the BBS (Datta et al., 2009). They also discovered that the usefulness of the individual BBS items varied across the phases of patients. The easier items on the BBS were more appropriate for the Phase I patients, whereas the more difficult items were more appropriate for the Phase III patients. This discrepancy suggests that the use of a sum of the BBS items may not be appropriate for all individuals with SCI. The researchers recommended the development of a new outcome measure, in which the items comprising the scale can be changed across the phase of recovery (Datta et al., 2009).

Wirz et al (2009) recently examined the concurrent validity and interrater reliability of the BBS in the chronic SCI population. In a study of 42 subjects with chronic SCI, the authors found strong correlations between the total BBS score and the scores on the Spinal Cord

Independence Measure II (SCIM) (r=.89, p<.001), the Falls Efficacy Scale (FES-I) (r=-.81, p<.001), the Walking Index for Spinal Cord Injury(WISCI) (r=.82, p<.001), gait speed as measured by the ten-meter walk test (r=.93, p<.001), and the ASIA motor score (r=.62, p<.001). These authors also found excellent interrater reliability of the BBS in this population, with an ICC= .953 for the total BBS score. Surprisingly, the authors did not find a correlation between the number of falls over a 5 month period and the BBS score. A ROC analysis was used to determine a BBS cutoff score for predicting falls. This analysis was unable to determine a cutoff score, because the area under the curve was .48, thus suggesting that the BBS "cannot discriminate beyond the chance of coincidence between participants who fall and those who do not." While these results are promising, one must note that all of the subjects tested were at least one year post injury, and therefore may have developed compensatory strategies to improve their balance and prevent falls. Furthermore, one third of the participants achieved a maximal score on the BBS, SCIM, FES-1, WISCI, and the ASIA motor score, indicating a high level of functional recovery, which may have skewed the results. As this study only established the concurrent validity of the BBS in the chronic SCI population, additional studies are needed to further establish the validity of this tool in all subjects with SCI.

Lemay and Nadeau (2009) recently established the concurrent validity of the BBS in the acute SCI population. This study examined the correlations between the BBS and the Walking Index for Spinal Cord Injury (WISCI II), the Spinal Cord Injury Functional Ambulation Index (SCI-FAI), the 10-meter walk test, and the Timed-Up and Go test (TUG), in a sample of 32 individuals with ASIA D incomplete SCI. The authors found that all of the walking tests were highly and significantly correlated with the BBS (0.714-0.816, $p<.01$). However, the authors also found a significant ceiling effect with the BBS in this population. Thus, the authors cautioned that the BBS should be used in conjunction with these walking tests when assessing a patient for appropriate assistive device use. This study was limited in its small sample size, acute population (mean time since injury was 77.2 days), and high-functioning sample. All subjects were able to ambulate at least 10 meters, with or without an assistive device, independently. This most likely resulted in the ceiling effect observed on the BBS. Thus, it is recommended that future studies, using a larger and more diverse population, be conducted to further examine the concurrent validity of the BBS in the incomplete SCI population.

Modified forms of the BBS.

Two studies have recently examined the psychometric properties of simplified versions of the BBS (Chou et al., 2006; C. Wang, Hsueh, Sheu, Yao, & Hsieh, 2004). Wang et al (2004) composed a 3-level version of the BBS by removing the third and fourth levels of scoring. Thus, a subject who could attempt to perform an item, but could not perform it perfectly, would be awarded the middle level of scoring for the item. The authors then compared the concurrent, convergent and predictive validity, and the responsiveness of this simplified BBS to the original scale in a two-part study. In the first part of the study, 77 patients with acute stroke were tested to establish the concurrent and convergent validity. In the second part of the study, 226 patients with acute stroke were tested to establish the predictive validity. The investigators found that the psychometric properties of the simplified version were equivalent to the original. The authors reasoned that a 3-level scale would be easier for clinicians and researchers to administer (C. Wang et al., 2004). However, further testing of this scale is needed to determine if the 3-level scale is sensitive enough to change in other neurological populations.

Recently, Chou et al (2006) argued that, based upon information reported by other researchers, the 14 items on the BBS required an increased amount of time to administer the test, and the 5-level items made scoring inconsistent. Thus, these authors developed eight shorter versions of the BBS, and tested each to determine which version had the psychometric properties closest to the original scale. The researchers found that the 7-item, 3-level version of the BBS had properties almost identical to the original (Chou et al., 2006). However, the simplified version did have a significant floor effect. The authors reasoned that this scale would take half the amount of time to administer than the original, as well reduce any inconsistencies in scoring. The simplified version also required fewer assessment tools (Chou et al., 2006). It is worth noting that while the 7-item test has been tested on 226 patients, they were all individuals with acute stroke. Further testing of the psychometric properties of this version is needed for use in other populations and at chronic time points post stroke.

The BBS has also been modified for use in the pediatric population. Franjoine et al (2003) modified the BBS by reordering several of the items, clarifying the directions, and reducing the time standards for several of the items. The authors reported a test-retest reliability ICC of 0.85 in children with normal development, and in children with mild to moderate motor impairment, the test-retest reliability ICC was 0.998, and the interrater reliability ICC was 0.997 (Franjoine, Gunther, & Taylor, 2003).

Current Uses of the Berg Balance Scale.

Current uses of the BBS in the elderly population.

The BBS has been used in a variety of ways in the elderly population. In a case study of a 101 year old female undergoing a frail-elderly exercise program, Gaub et al., (2004) found that the BBS score improved by 20 points. Li et al., (2004) in a study of 256 elderly subjects, found that individuals who participated in a Tai Chi exercise program had significantly greater improvement in BBS scores ($p<.001$) than those in the control group. Mihay et al., (2003, 2006) conducted two different studies on the effects of Tai Chi exercise on balance. In their 2003 quasi-experimental study of 35 elderly subjects, the authors found that participants in a Tai Chi experimental group performed significantly better ($p=.001$) on the BBS than the individuals in the control group, who received no intervention (L. Mihay et al., 2003).

However, subjects were not randomly assigned to the two groups, nor did they report the baseline BBS scores, so it is not known if the two groups differed at baseline (L. Mihay et al., 2003). In their 2006 study of 22 elderly subjects, Mihay et al., (2006) found no significant difference between improvements in the BBS scores in subjects participating in Tai Chi, versus those participating in a strength training program. Robinson et al., (2002) used the BBS to measure changes in fall risk in 25 elderly subjects undergoing a group physical therapy exercise program. These authors found that subjects classified as fallers had a significant improvement in BBS scores ($p<.05$) after the intervention, resulting in a decreased risk for falls (Robinson, Gordon, Wallentine, & Visio, 2002). Wolf et al., (2001), in a study of 94 elderly subjects, found that those who participated in an individualized exercise program significantly improved their BBS scores ($p<.001$). Hatch, Body, and Portney (2003) found that the BBS correlated with the Activity-specific Balance Confidence measure ($r= .752$), in a study of 50 elderly subjects. Thus, they reasoned that individuals with higher BBS scores had greater balance confidence (Hatch, Body, & Portney, 2003). The BBS has also been used in the elderly population to distinguish community-dwelling elderly female fallers from non-fallers (O'Brien, Pickles, & Culham, 1998). In 48 elderly females, O'Brien, Pickles and Culham (1998) found that fallers performed significantly worse than non-fallers. The BBS was also used to determine relationships between balance and other impairments. McAuly, Mihalko, and Rosengren (1997) determined the correlations between balance, as measured by the BBS, and fear of falling ($r= -.60$) in a study of 58 older adults. In a study of 20 elderly patients with a variety of diagnoses undergoing inpatient rehabilitation, Willems and Vandervoort (1996) found that all subjects had significant improvements in both BBS scores ($p<.001$) and gait speed ($p=.004$). These authors also found a strong correlation between BBS scores and gait speed ($r= .87$) (Willems & Vandervoort, 1996). In all of the studies mentioned above, the BBS appeared to be sensitive to changes in the study sample.

Current uses of the BBS in the CVA population.

The BBS has been used extensively in the CVA population. Thirteen studies used the BBS to evaluate the effectiveness of an intervention on improving the balance of individuals post stroke. Eng et al (2003) studied the effects of an 8-week community-based functional exercise program in a group of 25 individuals with chronic stroke. These authors found that subjects with an initial BBS below the median for the group had a significant improvement in BBS score after the intervention ($p<.001$) (Eng et al., 2003). Leroux, Pinet, and Nadeau (2006) also found significant improvements in BBS scores ($p<.05$) in 10 subjects with chronic CVA after an 8-week exercise program. However, these authors also reported that changes in the BBS score poorly correlated with increases in postural steadiness, as measured by COP displacements during functional tasks ($r= -.23$) (Leroux, Pinet, & Nadeau, 2006). This indicates that the BBS may assess different aspects of balance than are captured in COP measurements. In a 2005 study of 61 older adults with chronic CVA, Marigold et al., found that subjects in both an agility exercise group and a stretching/weight shifting group had significant changes in BBS scores ($p<.001$). In 2005, Mount, Bolton, Cesari, Guzzardo, and Tarsi published a case series describing a group balance class for 4 individuals with chronic CVA. After participating in this exercise class, the subjects improved between 1 and 5 points on the BBS. However, it should be noted that all subjects had an initial BBS score of >47/56, and therefore a ceiling effect may have occurred (Mount, Bolton, Cesari, Guzzardo, & Tarsi, 2005). In a 2006 study, Olawale and Ogunmakin found that an 8-week balance exercise training program significantly improved the BBS scores of 23 subjects with chronic CVA ($p<.05$). The authors found that while the BBS scores improved, the incidence of falls in the study's sample did not, as the subjects continued to experience falls (Olawale & Ogunmakin, 2006). In a single case study of 4 subjects with chronic CVA, Bastille and Body (2004) found that 8 weeks of a yoga-based exercise program only improved the BBS scores of 2 of the subjects. Pomeroy et al. (2001) found, in a study of 24 adults with chronic CVA, that training with weighted garments did not result in significant changes in BBS scores (p=.74). In 2001 Geiger, Allen, O'Keefe, and Hicks explored whether the addition of forceplate and visual feedback training would enhance other physical therapy interventions in 13 individuals with CVA undergoing an outpatient physical therapy program. These authors found that all subjects had a significant improvement in BBS score (p= .006), however, there was no significant difference between those subjects who had received the forceplate and visual feedback training and those who had not (Geiger, Allen, O'Keefe, & Hicks, 2001). These results are similar to those of Walker, Brouwer, & Culham (2000), who also found no significant difference in BBS scores between subjects who had received visual feedback training and those who had not, in a study of 46 subjects with acute CVA. Wang et al. (2005) studied the difference in BBS scores when the test was performed with, and without, an ankle foot orthosis (AFO). These authors found that in 42 subjects with acute CVA, and 61 subjects with chronic CVA, that the wearing of an AFO did not change the BBS score (Wang et al., 2005). Askim, Mørkved, and Indredavik (2006) evaluated the effect of an extended stroke unit service with early supported discharge on balance and walking speed in 62 patients with CVA undergoing inpatient rehabilitation. These investigators found that there were no differences in discharge BBS score or walking speed between patients on the extended service unit and the traditional stroke service unit (Askim, Mørkved, & Indredavik, 2006). Langhammer and Stanghelle (2003) studied the effects of 2 different CVA rehabilitation approaches, the Bobath technique and the Motor Relearning Program (MRP). The authors found that the mean BBS score for 61 subjects with chronic CVA was 19/56 for the MRP group and 20/56 for the Bobath group. The difference between these scores was non-significant; however the low scores indicated that all subjects were at increased risk of falling (Langhammer & Stanghelle, 2003). In a 2006 study, van Nes et al. examined the effects of a 6-week whole-body vibration intervention versus an exercise therapy with music intervention, on 53 patients in the post-acute phase of CVA. The authors found patients in both groups had increased BBS scores ($p<.01$), but there was no significant difference between the two groups (van Nes et al., 2006).

Three studies used the BBS as an impairment measure. Au-Yeung, Ng, and Lo (2003) performed a study to determine if ambulatory function is governed by motor impairment of limbs or balance ability in subjects with hemiplegia caused by stroke. Using the BBS to measure balance in 20 subjects with chronic CVA, the authors found that the BBS was able to distinguish between subjects who were able to ambulate independently, and those who required physical assistance during ambulation (Au-Yeung, Ng, & Lo, 2003). Eng et al (2002) studied the relationship between walk tests and measures of exertion (perceived and myocardial), in addition to impairment in 25 individuals with chronic stroke. These authors found that the BBS had good correlations with the six-minute walk test distance (r=.784), the 12-minute walk test distance (r=.798), and gait speed (r=.784) (Eng, Chu, Dawson, Kim, & Hepburn, 2002). Karatas, Çetin, Bayramoglu, and Dilek (2004) used the BBS to correlate whole-body balance with measures of trunk flexion and extension strength in 38 unihemispheric patients with CVA, and 40 healthy adults. The poor to moderate correlations of the BBS with the various trunk flexor and extensor torque values (r= .10-.64) suggested that trunk strength is only one small component of dynamic standing balance (Karatas, Çetin, Bayramoglu, & Dilek, 2004). The BBS has also been used as a measure of mobility. One study that utilized the BBS in this way studied the extent to which changes in functional mobility and balance are accompanied by changes in postural control in 27 patients currently undergoing rehabilitation (Garland et al., 2003). While the patients in this study exhibited a significant improvement in BBS scores ($p<.001$), there was no improvement in the activation of the paretic hamstring muscles. This suggested that these patients had regained functional balance through compensatory strategies (Garland et al., 2003). Another study which used the BBS as a mobility measure investigated the relationship between perceived health and mobility status in 13 people with chronic stroke (Au-Yeung et al., 2003). These authors found that the neither the BBS nor the TUG correlated with the Short Form-36, which was used to measure perceived health (Au-Yeung et al., 2003).

The BBS has also been used to establish correlations between balance and other factors relating to falls. Hellström, Lindmark, Wahlberg, and Meyer (2003) examined the correlation between the BBS and falls-related self-efficacy in 37 patients with acute CVA. Patients with low self-efficacy at discharge had less improvement in balance than patients with high self-efficacy (Hellström, Lindmark, Wahlberg, & Meyer, 2003). Mackintosh, Hill, Dodd, Goldie, and

Culham (2005) examined the relationship between the BBS and falls incidence and fall circumstances in 57 patients discharged from an acute rehabilitation hospital. Patients with lower BBS scores had higher incidence of falls, increased difficulty getting up after a fall, and restricted their mobility after the fall (Mackintosh, Hill, Dodd, Goldie, & Culham, 2005).

The BBS has been used to establish the validity of new outcome measures and was used to establish the construct validity of the Rivermead Mobility Index (RMI), an index used to measure the mobility of patients who have had a head injury or stroke (Hsieh, Hsueh, & Mao, 2000). In 38 patients with acute CVA undergoing inpatient rehabilitation, Hsieh, Hsueh, and Mao (2000) found that the BBS correlated well with the RMI at both admission (r= .81) and discharge (r= .89). Pal, Hale, and Skinner (2005) established the construct validity of the Activities-Specific Balance Confidence Scale (ABC), a tool originally designed for use with the elderly to measure the confidence while performing various activities of daily living, through moderate correlations with the BBS in 24 adults with CVA. Desrosiers, Rochette, and Corriveau (2005) used correlations with the BBS to establish the construct validity of the new Lower Extremity Motor Coordination Test

(LEMOCOT). In a validity study of 144 people with acute CVA, the BBS had a moderate correlation of .67 with the LEMOCOT (Desrosiers, Rochette, & Corriveau, 2005). Chern, Yang, and Wu (2006) attempted to establish construct validity of the Whole-Body Reaching (BBR) test through correlations with the BBS in a study of 23 subjects with acute CVA. However, the BBS only had fair correlations (r= .33), which were non-significant, with the WBR test (Chern et al., 2006). Tyson and DeSouza (2004) used correlations with the BBS to establish the criterion validity of a hierarchical series of functional balance tasks in a study of 48 subjects with acute CVA. Correlations with these tasks ranged from r=.26 during weight shifting, to r= .70 with standing forward reach, indicating that some of the functional tasks did not capture the same constructs of balance as the BBS (Tyson & DeSouza, 2004). The concurrent validity of the Trunk Control Test was established through correlations with the BBS (r=.755) in a study of 28 hemiparetic patients (Duarte et al., 2002).

The responsiveness of the BBS has also been compared to that of other measures. Hellström, Lindmark, and Fugl-meyer (2002) found that the responsiveness of the Falls-Efficacy Scale, Swedish version (FES-S), in detecting clinically meaningful changes over time was equal to the responsiveness of the BBS. Salbach et al. (2001) compared the responsiveness of the BBS and gait speed to the changes in ambulation ability in more severely affected individuals. In this study, the BBS was found to be more responsive than measures of gait speed.

Smith, Hembree, and Thompson (2004) compared the BBS to the Functional Reach test to determine the best clinical tool for measuring balance in individuals post stroke. If the data from the Functional Reach test, which is a shorter assessment, were strongly related to data from the BBS, then the clinician may choose to use the tool that takes the least amount of time to administer. The investigators found a strong correlation (r=0.78) between the two measures. However, the clinician must consider that the Functional Reach Test only assesses balance during one functional task, whereas the BBS contains items that can assess an individual's sensory integration, and thus provide a more detailed assessment of functional balance.

Four studies used the BBS to identify individuals with stroke who were at an increased risk of falling. Hyndman and Ashburn (2003) explored the relationship between attention, balance, function, and falls. In 48 adults with chronic CVA, the authors found that patients with better attention had higher BBS scores (p<.001) (Hyndman & Ashburn, 2003). Harris, Eng, Marigold, Tokuno, and Louis (2005) examined the relationship between balance, functional mobility, and falls in a study of 99 adults with chronic CVA and found that the BBS was unable to distinguish between individuals who had fallen once, and those who had never experienced a fall. However, only subjects who were primary manual wheelchair users were included, which may have confounded their results (Harris et al., 2005). These results are in contrast to those of Belgen, Beninato, Sullivan, and Narielwalla (2006), who aimed to determine how well the FES-S, the Timed-Up and Go test (TUG), and the BBS could distinguish between groups of subjects based on their history of falling. These investigators found that when a cutoff score of 52/56 was used, the BBS had a sensitivity of .91. They also found that subjects with a lower BBS score were more likely to experience multiple falls (p=.02) (Belgen, Beninato, Sullivan, & Narielwalla, 2006). Andersson, Kamwendo, Seiger Å, and Appelros (2006) used the BBS as one factor in describing the general characteristics of patients with stroke in inpatient rehabilitation who have a tendency to fall, and found that a combination of the BBS and the Stops Walking When Talking test yielded a sensitivity of .86 in identifying potential fallers.

The use of the BBS in a variety of studies shows that it is the outcome measure of choice when assessing balance in the CVA population. While the BBS was appropriate for use, not all investigators found the BBS to be responsive enough to changes, and did not correlate well with all other outcome measures.

Current uses of the BBS in the MS population.

Three studies in individuals with multiple sclerosis (MS) have used the BBS to assess the effects of an intervention. A pilot study used the BBS as a secondary assessment when comparing two physical therapy approaches to improve walking in patients with gait disturbance due to MS (Lord, Wade, & Halligan, 1998). These authors found that both groups showed a significant improvement in BBS scores ($p<.001$) (Lord, Wade, & Halligan, 1998). Hale et al (2003) used the BBS to measure changes in balance after an eight-week program of aerobic, stretching, strengthening and balancing exercises in four people with MS. These authors found substantial changes in BBS scores after this intervention ($p<.01$) (Hale, Schou, Piggot, Littmann, & Tumilty, 2003). Most recently, Smedal et al (2006) used the BBS to measure changes in balance after physical therapy based on the Bobath concept was applied to 2 patients with MS who had balance and gait problems.

These authors found 1 to 2 point improvements in BBS scores for both patients, although the changes were not significant (Smedal et al., 2006). Although Smedal et al. (2006) did not report significant changes in BBS scores, the other 2 studies mentioned above (Hale et al., 2003; Lord et al., 1998) show that the BBS can be responsive to changes in balance in the MS population.

Current uses of the BBS in the PD population.

In the Parkinson's disease (PD) population, only one study used the BBS to assess change after an intervention. Toole, Maitland, Warren, Hubmann, and Panton (2005) measured the effects of loaded and unloaded treadmill walking on balance, gait, fall risk and daily function in patients with PD. The authors found that all patients who participated in treadmill walking, whether loaded or unloaded, made significant progress on all outcome measures. Therefore, the authors reasoned that the amount of weight bearing while walking on the treadmill had no effect on the outcome measures (Toole, Maitland, Warren, Hubmann, & Panton, 2005).

Current uses of the BBS in the SCI population.

To date there is very little literature that has examined the psychometric properties of the BBS in the SCI population. A review of 65 articles on SCI, walking training, and balance revealed that only 3 articles used the BBS as an outcome measure (Behrman & Harkema, 2000; Dobkin, Apple, Barbeau, & et al., 2006; Musselman, Fouad, Misiaszek, & Yang, 2009). One article, a series of case studies, used the BBS to measure balance outcomes on one of the four subjects (Behrman & Harkema, 2000). The authors of this article (Behrman & Harkema, 2000) noted concerns that the BBS had not yet been validated for use in the SCI population. The second study, a multi-center randomized controlled trial on locomotor training in patients with acute incomplete spinal cord injury, used the BBS as a secondary outcome measure (Dobkin et al., 2006). The BBS results were not published in that study. The third study, a case series of four individuals with incomplete SCI, used the BBS to measure balance outcomes on all four of the subjects (Musselman, Fouad, Misiaszek, & Yang, 2009). The authors found minor changes in balance in all subjects, reporting change scores between 0 and 10 points on the BBS, across the 4 subjects. The authors argued that a floor effect may have occurred as the BBS does not permit the use of assistive devices, and all subjects required an assistive device to be able to stand (Musselman et al., 2009). One must remember that only the concurrent validity has been established for the BBS in the chronic SCI population, and therefore further studies are needed to ensure that the BBS is truly measuring balance in this population. Also, as the minimal detectable change (MDC) has not been established for this population, we do not know what amount of change needed to establish significance in individuals with SCI. Furthermore, the BBS only has one item that measures sitting balance, and therefore individuals with SCI who are unable to stand and complete any of the standing tasks, will experience a floor effect, as they will not be able to improve their score beyond improvement in that one item. Individuals who are ambulatory without an assistive device may experience a ceiling effect on the BBS. While these individuals may have high-level balance deficits, especially exhibited during ambulation, the BBS may not pick up these deficits due to the lack of a gait component. Therefore, the BBS is not able to capture changes in balance across the continuum of recovery in individuals with SCI.

Development of the Tinetti Performance-Oriented Mobility Assessment (POMA)

The initial purpose of the POMA was to assess fall risk in the elderly (Tinetti ME, William TF, Mayewski R, 1986; Tinetti, 1986). It was designed to combine diagnosis, therapy, and assessment of interventions (Tinetti, 1986). The POMA consists of two indexes: one which measures balance; the other assesses gait. In the original version of the scale, each of the indices contained 8 items, with a total possible score of 28. These indices were developed by reviewing the works of bioengineers, orthopedists, neurologists, rheumatologists, and physical therapists. The process for selection from each of these works, as well as the development of the scoring criteria, were not discussed (Tinetti ME, William TF, Mayewski R, 1986; Tinetti, 1986). The authors stated that they fashioned the scoring of the scale to be vague, in order to allow mobility to be evaluated in a practical manner (Tinetti ME, William TF, Mayewski R, 1986).

In early works on the POMA, the scale creators reported basic psychometric properties. Pre-testing, by two different raters, of the original scale on 15 ambulatory elderly residents of a long-term care facility revealed an interrater reliability of 90% (Tinetti ME, William TF, Mayewski R, 1986; Tinetti, 1986). In a later project, 5 more balance maneuvers were added to the scale, and the authors reported an interrater reliability of 85% (Tinetti, 1986).

The researchers attempted to validate the POMA through correlations with several physical impairments. They found that the POMA had good to fair correlations with lower extremity strength ($r=.55$), active back extension ($r=.45$), and active neck extension ($r=.37$) (Tinetti ME, William TF, Mayewski R, 1986). The investigators also discovered that the POMA was able to discriminate between fallers and non-fallers. Although sensitivity and specificity levels were not reported, the authors did report significant differences ($p<.0001$) in mean scores between recurrent fallers (14 ±6) and non-fallers (21 ± 4) (Tinetti ME, William TF, Mayewski R, 1986).

Psychometric Properties of the POMA

Limited research has been published on the psychometric properties of the POMA. It is difficult to compare the literature that does exist, because some studies used only the balance or the gait index, while others used versions that incorporated balance or gait maneuvers other than the original 8 described. Finally, very few of these studies stated the exact version of the POMA that was used.

The psychometric properties of the POMA in the elderly population.

The test-retest and interrater reliability of the POMA in ambulatory, elderly individuals with few established neurological conditions have been reported (Cipriany-Dacko et al., 1997; Lin et al., 2004; Mecagni, Smith, Roberts, & O'Sullivan, 2000; Tinetti et al., 1993). Both the intrarater and interrater reliability ranged between 0.93 and 0.99 for both the original and longer versions of the POMA in well-elderly subjects (Lin et al., 2004; Mecagni et al., 2000; Tinetti et al., 1993). Cipriany-Dacko et al (1997) investigated the interrater reliability of the original version of the balance index. Nine raters, including experienced and novice clinicians, as well as physical therapy students, assessed the balance of 29 elderly subjects. Interrater reliability was computed through Kappa coefficients, ranging from 0.40 to 0.75 among all nine raters and indicated that only minimal training is needed to reliably apply the scale in an elderly population (Cipriany-Dacko, Innerst, Johannsen, & Rude, 1997). Faber, Bosscher, and van Wieringen (2006) examined the reliability of the original balance and gait indexes separately, as well as combined. These researchers found the test-retest reliability to range from 0.72 to 0.86, and an interrater ICC range of 0.80 to 0.93.

Concurrent validity of the POMA has been investigated in a prospective study of 167 older adults. Cho, Scarpace, and Alexander (2004) found moderate to good correlations of the original POMA with tandem stance time, single leg stance time, maximum step length, and the Activity-specific Balance Confidence Scale. A Spearman correlation of 0.79 between the original POMA and gait impairment scores was found in a study of 59 elderly subjects (Baloh, Ying, & Jacobson, 2003). Lin et al (2004) reported moderate to strong correlations between a longer version of the POMA and the functional reach test, gait speed and the Timed-Up and Go (TUG) test. Faber et al (2006) found that the balance index of the POMA, and the total POMA scale had stronger concurrent validity than the POMA gait index alone. The POMA balance index and the total POMA scale correlated highly with the TUG, gait speed, and the FICSIT-4 balance scale (Faber et al., 2006).

Discriminant validity was only examined in two studies (Faber et al., 2006; Lin et al., 2004). Lin et al (2004) found that the POMA had a high discrimination power for falls history, assistive device use, and disability with activities of daily living. Faber et al (2006) also reported that the POMA was able to discriminate between assistive device use in a sample of 245 elderly persons.

Faber et al (2006) assessed the validity of the POMA in predicting the number of falls of each elderly subject over a 10-month span. These researchers determined that the predictive validity was not satisfactory, and reported a sensitivity of 64% and a specificity of 66%. These findings were consistent with those of Verghese, Buschke, Viola, Katz, Hall, Kuslansky, et al., (2002), who found a sensitivity of 61.5% and a specificity of 69.5% with a cut-off score of 10, out of 16, on the POMA balance index. These researchers also determined the positive predictive value of the POMA balance in predicting falls to be 36.4%. Therefore, using this test alone, an examiner might miss 64% of fallers. Furthermore, Raiche, Herbert, Prince, and Corriveau (2000) discovered low predictive validity. The researchers reported a sensitivity of 70% and specificity of 52% in a prospective study of 225 subjects. It is worth noting that these authors used an extended version of the POMA, with 14 balance and 10 gait items (Raiche et al., 2000). These results are in disagreement with those reported by Chiu, Au-Yeung, and Lo (2003). These investigators reported finding 95% sensitivity and specificity in predicting multiple fallers using a longer version of the POMA balance index.

The validity of the POMA gait index in predicting death or nursing home placement over a 22-month span was assessed (Reuben, Siu, & Kimpau, 1992). In a prospective study of 149 elderly subjects, Reuben, Siu, and Kimpau (1992) found that this measure was a significant predictor of death or nursing home placement ($r=.19$; $p<.05$).

The ability of the POMA to detect changes in balance.

Faber et al (2006) determined the minimal detectable change score to be 5 points on the original total POMA. Shore, deLateur, Kuhlemeier, Imteyez, Rose, and Williams (2005) compared the responsiveness of the POMA gait index to the GAITRite walkway in assessing change in the ambulatory performance of patients with normal pressure hydrocephalus before and after shunt surgery. The authors found that there were discrepancies in the amount of change noted by the POMA and the GAITRite. They discovered that there was a significant ceiling effect with the POMA, resulting in decreased sensitivity to change. They recommended that the POMA be used as a screening tool for identifying gross gait deviations, however suggested that the GAITRite be used to measure fine discriminations (Shore et al., 2005).

The psychometric properties of the POMA in the PD population

Kegelmeyer, Kloos, Thomas, and Kostyk (2007) examined the psychometric properties of the POMA as a tool for screening risk of falling in the PD population. In a prospective study of 30 subjects with PD, the intrarater reliability ranged between 0.69 and 0.88, and the interrater reliability ranged from 0.80-0.94. The authors used these same patients to establish concurrent criterion validity through moderate to good correlations (r= -0.40 to –0.45) with the Unified Parkinson's Disease Rating Scale (UPDRS), and with comfortable gait speed (r= 0.50 to 0.53). Finally, the authors conducted a retrospective study of 126 patients' records to identify the sensitivity and specificity of the POMA in detecting fall risk for patients with PD. The investigators found that the sensitivity was 76%, and the specificity was 66%. A cutoff score of 20, out of 28, was found to optimize the sensitivity and specificity ratios in order to best identify fallers (Kegelmeyer et al., 2007).

Behrman, Light, and Miller (2002) assessed the sensitivity to change of the POMA gait index in patients with Parkinson's disease

(PD). These researchers found that the POMA could distinguish between subjects with and without PD, but could not detect meaningful changes in the gait patterns of patients with PD. These authors agreed with Shore et al (2005), and recommended the POMA for use as a gross screen, but urged caution when using this tool to assess the effect of an intervention in patients with PD.

Current Uses of the POMA

Current uses of the POMA in the elderly population.

Seven recent studies have used the POMA to assess changes in balance after an intervention in the elderly population. Five of these studies investigated the effects of different types of exercise on the balance of elderly individuals (Bruin & Murer, 2007; Sauvage et al., 1992; L. Taylor et al., 2003; L. F. Taylor et al., 2003; Urbscheit & Wiegand, 2001). Significant changes in POMA scores were observed after interventions in all of these studies, with the exception of Urbscheit and Wiegand (2001). Urbscheit and Wiegand (2001) found that the initial score influenced the amount of change in the POMA score following an 8-week exercise program in elderly individuals. The POMA was also used to examine the effect of balance training under different visual conditions (Huang, Burgess, Weber, & Greenwald, 2006), and to determine the effects of whole body vibration (Bruyere et al., 2005). Huang, Burgess, Weber, & Greenwald (2006) found that there was a significant interaction between the POMA and vision ($p<.01$), in a study of 89 adults with balance impairments. In a study by Bruyere et al. (2005), a significant change in POMA scores ($p<.001$) was seen in elderly nursing home residents who underwent a whole body vibration intervention as compared to a control group. Thus, the POMA has been shown to be sensitive to changes in balance in the elderly population.

One study recently tested the relationship between the POMA and a fear of falling index. Manning, Neistadt, and Parker (1997) found that there was no significant correlation between the POMA and a fear of falling index.

One study used the balance subscale of the POMA to establish the concurrent validity of the Fast Evaluation of Mobility, Balance, and Fear (FEMBAF) in the elderly population (Di Fabio & Seay, 1997). The POMA was found to correlate significantly with four of the six components of the FEMBAF ($r= 58-.91$).

Two studies used the POMA to examine fall risk in elderly individuals. Trueblood, Hodson-Chennault, McCubbin, and

Youngclarke (2001) determined the sensitivity and specificity of performance and impairment-based tests in predicting falls. These authors found that an impairment-based test, the Limits of Stability test, was most accurate at predicting fallers, while the POMA, a function-based test, was most accurate at predicting non-fallers (Trueblood, Hodson-Chennault, McCubbin, & Youngclarke, 2001). Conner-Kerr and Templeton (2002) evaluated chronic risk factors for falls in elderly individuals with type II diabetes. Impaired balance, as measured by the POMA, was found to be a risk factor for falls in 40% of the subjects from an urban day care center (Conner-Kerr & Templeton, 2002).

Current uses of the POMA in the CVA population.

Two investigations used the POMA to evaluate changes in gait after an intervention in the CVA population. In a 2002 case report, Miller, Quinn, and Seddon reported the feasibility for using a body weight support system for overground ambulation as well as the measurement of functional changes in 2 patients with chronic stroke before and after body weight support training. In this case report, the authors reported increases in 2 points on the POMA gait index for both patients, although it was not stated whether the results were significant (Miller, Quinn, & Seddon, 2002). Daly, Roenigk, Holcomb, Rogers, Butler, Gansen, et al. (2006) tested the response to coordination exercise, overground gait training, and body weight supported treadmill training, both with and without functional neuromuscular stimulation using intramuscular electrodes in 32 patients with stroke. These authors found that there was a statistically significant improvement in POMA-gait scores ($p=.003$) for the patients receiving functional neuromuscular stimulation, versus the control group (Daly et al., 2006).

Only one study has used the POMA to discriminate between non-fallers, fallers, and repeat fallers. Soyuer and Öztürk (2007) used the POMA in their study on the effects of spasticity, sensory impairment, and type of walking aid on falls in community dwellers with chronic stroke. These authors found that the POMA was able to significantly distinguish between fallers, non-fallers, and repeat fallers ($p<.001$) (Soyuer & Öztürk, 2007).

Corriveau et al (2004) used the POMA as a clinical balance measure in their comparison of clinical and biomechanical measures of balance in elderly stroke patients with those of healthy elderly people. They found that the POMA had correlations of -.57 and -.58 with anterior/posterior and medial/lateral center of pressure amplitude measurements (Corriveau, Hébert, Raîche, & Prince, 2004). Bainbridge, Davie, and Haddaway (2006) used the POMA as a mobility measure in their study of bone loss in individuals with stroke. These authors correlated the rate of bone loss to measures of function and mobility. They found that low scores on the POMA were indicative of increased bone loss (Bainbridge, Davie, & Haddaway, 2006).

Current uses of the POMA in the MS population.

The POMA has been used to assess changes of balance in individuals with MS. In a case report, Fell (2000) reviewed current theories of mental imagery and practice, and described the implementation of such a program for an individual with MS. While the subject in the study only demonstrated an improvement of 1 point on the POMA, a ceiling effect may have occurred, as the subject initially scored a 14/16 on this assessment (Fell, 2000).

Current uses of the POMA in the SCI population.

To date there is no literature that has examined the psychometric properties of the POMA in the SCI population. Furthermore, a review of the literature did not reveal any studies that have used the POMA as an outcome measure in the SCI population.

The POMA is often used in a clinical setting because it has a gait component, and is therefore seen as being able to assess balance during ambulation. While this may be true for some higher-level balance deficits, the items that assess balance during gait are basic, and reflect the quality of the gait more than the patient's ability to maintain balance during different ambulation conditions. The POMA still faces the same challenges for use in the SCI population as does the BBS. As the POMA has not been validated for use in individuals with SCI, we cannot be certain that the phenomenon of balance is truly being measured. Also, as the MDC has not been established for this population, we do not know what a significant amount of change is for individuals with SCI. Furthermore, the POMA, like the BBS, only has one item that measures sitting balance, and consequently also may exhibit a floor effect. Therefore, the POMA also is not able to assess balance across the continuum of recovery in individuals with SCI.

The Development of the Modified Functional Reach Test

The purpose of the Modified Functional Reach Test (MFRT) was to provide an accurate tool for measuring sitting balance in the SCI population (Lynch, Leahy, & Barker, 1998). Lynch et al (1998) adopted the MFRT from the Functional Reach test, originally designed by

Duncan, Weiner, Chander, and Studenski (1990). While the original FRT was performed with the subject standing, the MFRT required that the subject sit with hips, knees and ankles at 90°, with one upper extremity raised to 90° shoulder flexion. This posture was adopted in order to ensure standardization and accuracy of measurement for all subjects (Lynch et al., 1998). The distance that the subject could reach forward was measured with a yardstick.

The Psychometric Properties of the MFRT

The original work on the MFRT used 30 male subjects with motor complete spinal cord injuries, who were able to maintain 90° of shoulder flexion and reported a test-retest reliability ICC ranging from 0.85 to .94 (Lynch et al., 1998). However, the ICC values may have been skewed as only one tester was used for the entire study. The authors found that the MFRT was able to distinguish between subjects with high-level lesions (C5-6, T1-4) and subjects with low paraplegia (T10-12). The MFRT was unable to distinguish between subjects with tetraplegia and high paraplegia. The authors discussed that the MFRT had good face validity, because it adequately challenged the subjects in each group (Lynch et al., 1998).

Another study by Adegoke, Ogwumike, and Olatemiju (2002) attempted to further establish the reliability of the MFRT, using 20 non-standing SCI subjects, who met the same inclusion criteria as in Lynch et al (1998). Only one rater was used and the ICC values for test-retest reliability ranged from 0.981 to 0.992 (Adegoke, Ogwumike, & Olatemiju, 2002). Unlike Lynch et al (1998), the authors did not detect any significant differences between patients with high-level lesions and patients with low paraplegia (Adegoke et al., 2002).

Current Uses of the MFRT

One study used the MFRT for comparison with the results of virtual reality testing (Kizony, Raz, Katz, Weingaden, & Weiss, 2005). The investigators modified the testing protocol by allowing subjects to support themselves with 1 upper extremity, while reaching with the other upper extremity. Additional measures of reaching right and left while facing the measurement tape were added (Kizony, Raz, Katz, Weingaden, & Weiss, 2005). The reliability and validity of any of these measures was not reported.

There is an obvious dearth of information on the psychometric properties of the MFRT. Further studies are needed to assess the interrater reliability, as well as establish construct, concurrent and predictive validity of this measure.

The MFRT is currently used in the clinical setting to capture changes in sitting balance in patients who are unable to stand, and who therefore cannot complete the majority of the items on the BBS and the POMA. However, the MFRT is limited in the amount of information that it provides, as it only measures how far a patient can reach in one direction. While this does give therapists and researchers some idea of how the patient's sitting balance is progressing, it is not a complete assessment of the patient's functional abilities in sitting. Therefore, the MFRT provides us with a good foundation for the assessment of sitting balance, but more items are needed to gain a more comprehensive picture of the changes in sitting balance of patients who are unable to stand.

The literature review presented demonstrates that the Berg Balance Scale, the Performance-oriented Mobility Assessment, and the Modified Functional Reach Test all have limitations when applied to the individuals with SCI. Therefore, a new outcome measure is needed to assess balance in this population. The new outcome measure should incorporate new items, which will assess balance in both the early and late stages of recovery, with the most relevant items of the BBS and the POMA, as determined through a principle component analysis. The new instrument should have several items that provide a more comprehensive assessment of sitting balance, as well as an added gait component that integrates higher-level items, such as walking while carrying an object, stepping over an object while walking, or turning around 180 degrees to quickly reverse directions. One procedure for developing a new outcome measure will be described below.

Outcome Measure Development

The Delphi Technique

One of the first steps in developing a new outcome measure is to generate items for the tool. One method of generating new items is obtaining a consensus amongst a group of experts in the field, regarding the importance, wording, and scoring of the items. A Delphi technique is a process that is often used to obtain a reliable consensus amongst a group of experts, using a series of questionnaires. In this technique, the experts anonymously provide their opinions on a first-round survey. The responses are summarized and reported back to the experts in the next round. This continues until a consensus is reached, or the response rate diminishes (Hasson, 2000).

There are several characteristics that distinguish the Delphi technique (Hsu & Sanford, 2007; McKenna, 1994). These characteristics include 1) the use of experts, 2) the preservation of anonymity, 3) the use of controlled feedback, 4) the ability to use statistical analysis on the responses, 5) the participants do not meet face to face, 6) the use of two or more rounds of sequential questionnaires (Hsu & Sanford, 2007; McKenna, 1994). The use of the Delphi to maintain anonymity and control feedback prevents friction amongst the respondents and allows for the group to be guided in generating a consensus (McKenna, 1994). This differs from the typical group experiences of expert panels or round table discussions, where the experts are identified and meet face-to-face. In these experiences, discussions can become heated, and some members may influence the group's decisions (Streveler, Olds, & Miller, 2003). Thus, in the Delphi technique, by maintaining anonymity and controlling the feedback, the investigator can ensure that each individual's opinion is weighted equally (McKenna, 1994, Streveler et al., 2003). Furthermore, as the Delphi can be conducted via questionnaires distributed by mail or electronically, experts across a wide geographical area can be surveyed (Skulmoski, Hartman, & Krahn, 2007).

The Delphi technique has many uses in research as demonstrated by Ditunno et al. (2000), who used a modified Delphi technique when developing the Walking Index for Spinal Cord Injury (WISCI). During the Delphi study, three professionals from eight internationally recognized spinal cord treatment centers were asked to rank the items on the WISCI. The Kendall coefficient of concordance and Spearman's rank correlation coefficient to determine the amount of agreement amongst the experts were calculated. The experts were provided with these results, and were asked to come to a consensus amongst each of the clinics for any areas that exhibited a discrepancy. After the second round of consultation, a consensus was reached.

The Delphi technique has several limitations. It is often criticized for appearing to force a consensus, and for not allowing participants to fully explain their opinions (Hasson, Keeney & McKenna, 2000). Furthermore, one must keep in mind that because a consensus has been reached, it does not mean that this consensus is the correct answer (Hasson, Keeney & McKenna, 2000). Therefore, while a Delphi technique can be employed to generate a new outcome measure, the new scale must then undergo extensive testing to ensure that it has the correct scope and depth required.

Rasch Analysis

The Rasch analysis is a statistical model that can estimate the person "ability" and item "difficulty" of a measurement tool by comparing the responses of individuals to the entire sample (Duncan, Bode, Lai, Perera, & Glycine Antagonist in Neuroprotection Americas Investigators, 2003). This model provides a method to analyze and improve a rating scale (Linacre, 1999). The greatest benefit of using a Rasch analysis is it converts ordinal measures into interval scales (Bond & Fox, 2001; Chang & Chan, 1995; Duncan et al., 2003; Tsuji, Meigen, Sonoda, Domen, & Chino, 2000; G. Williams, Robertson, Greenwood, Goldie, & Morris, 2005; Wright & Masters, 1982). The advantage to developing an interval-level scale is parametric statistics can then be used with this scale, which strengthens the results of the study (Portney & Watkins, 2000).

A Rasch analysis is used to test three specific properties of a rating scale, including targeting, item difficulty, and person separation. Targeting is the range of difficulty of the items. Testing the item difficulty may reveal any clusters of items, or items that appear to have the same level of difficulty, and are redundant. Gaps in the difficulty level of the scale can be examined, resulting in the possible need to modify or create new items (Bond & Fox, 2001). Furthermore, using a Rasch analysis to test item difficulty allows for the items to be placed in order of increasing difficulty (Bond & Fox, 2001). The sample population can also be placed in order of increasing ability, by calculating the person separation index. This index allows the researcher to identify distinct functional levels in a sample, and also shows if the outcome measure has an inherent floor or ceiling effect (Duncan et al., 2003, Bond & Fox, 2001).

Rasch analysis has been used by many researchers in the development of new health measurement scales. Duncan et al (2003) used a Rasch analysis to evaluate different psychometric properties of the Stroke Impact Scale (SIS), including targeting, item difficulty, and separation. Targeting measures whether the items are of an appropriate level of difficulty for the sample population. Item difficulty refers to the order of the items, from least to most difficult. Separation refers to whether the items are able to distinguish distinct functional levels in a sample (Duncan et al., 2003). In this study, Duncan et al.

(2003) used a sample of 696 subjects with stroke participating in the Glycine Antagonist in Neuroprotection (GAIN) Americas randomized clinical trial. The version of the SIS that was tested in this study consisted of 64 items across 8 domains. The authors found that the targeting of the SIS was appropriate, as it was able to capture a large range of difficulties, and the order of the items did progress from less difficult to more difficult in a clinically meaningful manner. The authors found that the separation ability of the physical domains of the SIS was adequate, as these domains were able to distinguish more than 4 levels of functioning. However, the communication, memory, and emotion domains were only able to distinguish 2 to 3 levels of functioning, indicating that these domains might only be useful in very low functioning patients. Thus, the results of the Rasch analysis guided these authors on how to revise and improve the SIS (Duncan et al., 2003).

Tsuji et al. (2000) used a Rasch analysis to examine the item difficulty of the Stroke Impairment Assessment Set (SIAS) in a sample of 190 patients with stroke undergoing inpatient rehabilitation. The SIAS was developed to assess various impairments in patients with hemiplegia (Tsuji et al., 2000). The analysis revealed a good fit of most items, with the exception of 4 items. Based upon these findings, the authors stated that they would consider dividing the instrument into subscales to improve the fit of these items. The authors also determined that the item difficulty patterns were the same at admission and discharge, with the exception of 3 items. Thus, the authors concluded that they needed to combine several items in order to improve the quality of the SIAS (Tsuji et al., 2000).

The Rasch analysis has been used to establish the content validity and discriminability of the High-level Mobility Assessment Tool (HiMAT) (G. P. Williams, Robertson, Greenwood, Goldie, & Morris, 2005). The HiMAT is an outcome measure that was developed to assess high-level mobility in individuals with traumatic brain injury. The initial version of the HiMAT consisted of 28 items. In a sample of 103 patients with traumatic brain injury, Williams et al. (2005) used a three-step process to establish content validity. This process began by testing the internal consistency of the items by calculating Cronbach's alpha. The authors then utilized principal axis factoring to find the linear correlations between the items. The principal axis factoring resulted in the removal of 8 items from the scale, and the separation of the stair item into 2 separate items. Finally, a Rasch analysis was used to test the unidimensionality of the remaining 22 items to identify any misfitting items. Two misfitting items were excluded from the final version of the scale, resulting in a total of 20 items. A Rasch analysis was further utilized to test the discriminability of the HiMAT. The authors used the item estimates from the Rasch analysis to identify items that were clustered at the same level of difficulty. The authors then eliminated several redundant items based upon these clusters, by removing the items that were considered to be more difficult to test. Thus, the Rasch analysis resulted in the development of the final version of the HiMat (Williams et al., 2005).

In 2004, Kornetti et al. conducted a Rasch analysis of the Berg Balance Scale to determine the effectiveness of the scoring criteria for each item. The authors found that when underutilized scoring categories for each item were condensed, and the rating scale was re-scored, changes in the item difficulty order became apparent. Prior to the collapsing of rating scale categories, the most difficult item on the scale was "Standing on one leg," with an item difficulty of 3 logits. This resulted in a ceiling effect of the BBS, as the subject with the highest balance ability had an ability level of 6 logits. After collapsing the rating scale categories, and re-scoring the scale, the most difficult item was "Tandem stance," with an item difficulty of 6 logits. This resulted in eliminating the ceiling effect of the BBS. These changes also resulted in an improved spread of item difficulties, and the items were more evenly distributed across the sample of persons tested. In this study, the Rasch analysis was used to improve an already-established outcome measure.

Based upon the literature, it appears that a Rasch analysis is an appropriate intermediate step to develop a clinical outcome measure. A Rasch analysis can be utilized to test the targeting, item difficulty, and separation of a new tool. Based upon the findings from this analysis, the authors of the tool would be able to refine the new scale.

Chapter III

METHODS

Introduction

The Activity-based Balance Level Evaluation (ABLE scale) was developed in collaboration with the Balance Committee of the NeuroRecovery Network. In 2004, the Christopher and Dana Reeve Foundation (CDRF) and the Centers for Disease Control and Prevention (CDC) established the NeuroRecovery Network (NRN). The mission of the NRN is to "provide support for the development of specialized centers that provide standardized activity-based therapy based on current scientific and clinical evidence for people with spinal cord injury and other selected neurological disorders. (The NeuroRecovery Network, 2009)" The NRN currently consists of 7 nation-wide centers that provide locomotor training for patients with incomplete spinal cord injury. The Balance Committee of the NRN has been charged with the task of developing a new outcome measure that is sensitive across a wide range of balance level ability in the spinal cord injury population, can be implemented with minimal equipment, and can be completed in a timely manner. The author of this study is the chairperson of the Balance Committee.

Subjects

A total of 104 subjects with SCI were recruited for this study. This sample size was chosen based upon the work of Wang et al. (2005) (W. Wang & Chen, 2005), who recommended the use of 100 subjects when using a Rasch analysis to analyze a rating scale containing 20 items. Subjects were recruited from the inpatient and outpatient settings of Magee Rehabilitation Hospital in Philadelphia, PA, Shepherd Center in Atlanta, GA, Kessler Research Center in West Orange, NJ, and Frazier Rehabilitation Institute in Louisville, KY. Inclusion criteria for this study specified that subjects were at least 16 years of age and had a traumatic or non-progressive, complete or incomplete spinal cord injury. Exclusion criteria included: inability to follow 2-step commands, need for a spinal stabilization device, spinal precautions which limit the ability to bend or rotate in the thoracic or lumbar spine, inability to tolerate upright supported sitting for at least 1 minute.

Design and Variables

This exploratory research study utilized a methodological research design to test the initial psychometric properties of a new balance outcome measure for the SCI population, the ABLE scale. The properties tested were item difficulty, and person ability. Item difficulty is defined by Bond & Fox (2001) as "an estimate of an item's underlying difficulty calculated from the total number of persons in an appropriate sample who succeeded on that item." Person ability, as defined by Bond & Fox (2001), is "an estimate of a person's underlying ability based on that person's performance on a set of items that measure a single trait."

Measurement Tools

Demographic data and the Activity-based Balance Level Evaluation (ABLE scale) will be the two main measurements collected.

Demographic data.

The ABLE scale captured demographic information from each subject, using demographic items that were designed by the author, and included age, gender, date of injury, and questions regarding the type and severity of the spinal cord injury. Additional self-report items aimed to determine a general functional level for the subject (Appendix A).

The Activity-based Balance Level Evaluation.

The Activity-based Balance Level Evaluation (ABLE scale) was developed in collaboration with the Balance Committee of the NRN. To develop items for the ABLE scale, the primary author employed a Delphi technique. During this process, experts in SCI rehabilitation from the United States completed a series of surveys in which they graded the importance of each item, the scoring of the item, and the wording of the item (Appendix B). This process resulted in a total of 28 items, which assess balance across the three domains of function: sitting, standing, and walking (Appendix A). Each item is scored on a 5-point ordinal scale. After the Delphi technique was completed, a brief pilot study on two subjects with incomplete SCI was done to test the feasibility of administering the ABLE scale in a physical therapy clinic (Appendix C). During this pilot study, two physical therapists separately administered the ABLE scale on two subjects with incomplete SCI, who were at different ends of the spectrum of recovery. The physical therapists commented on the time to complete the ABLE scale, the clarity of the items and scoring, and the overall flow of the outcome measure. Interrater reliability between the two physical therapists was calculated to range from .833 on the subject who could stand and walk without an assistive device, to .944 on the subject who could only sit (Appendix C). Although the Delphi technique resulted in a scale with content validity, through the consensus reached by experts, and the pilot testing on subjects with incomplete SCI established the feasibility of administering this outcome measure, the ABLE scale required additional testing in order to further refine the scale. Additional testing included the assessment of redundant items and the assessment of gaps in the level of difficulty across the scale, which would make it necessary to either remove, modify, or add new items.

Procedure

Approval for the study was granted from the IRB at Magee Rehabilitation Hospital, Shepherd Center, Kessler Research Center, Frazier Rehabilitation Institute and Seton Hall University. To ensure standardization of the scoring and administration of the ABLE scale across the data collection centers, the primary investigator provided an instructional session at the NRN's annual National Summit at Frazier Rehabilitation in Louisville, KY, in January 2009. The electronic presentation used during this session was also provided to any physical therapists involved in the data collection, who were unable to attend the onsite instructional session. The primary investigator also responded to any concerns the therapists had regarding the administration and scoring of the ABLE scale via phone call or email.

All participants were asked for their consent to participate in this study by the primary investigator, or one of the designated physical therapists at the 4 data collection sites. These patients were then tested on the ABLE scale by the primary investigator (at Magee Rehabilitation Hospital), or one of the designated physical therapists at the 4 data collection sites (Magee Rehabilitation Hospital, Shepherd Center, Frazier Rehabilitation Institute and Kessler Research Center) outside of their scheduled physical therapy treatment time. These data were entered into an Excel spreadsheet, and sent to the primary investigator by the data collector. All data were saved on a thumb drive which was kept in a locked drawer in the primary investigator's desk at Magee Rehabilitation Hospital.

Data analysis.

Descriptive statistics were used to analyze the demographic data. A Rasch analysis was completed on the ABLE scale to evaluate the following psychometric characteristics: targeting, item difficulty and person separation. While the ABLE scale consists of 28 items, 2 of the items (item 3 and item 21) were separated into a right and left component for the purposes of the Rasch analysis. Thus, a total of 30 items underwent the analysis.

It is important that an outcome measure assesses subjects across the full spectrum of ability. Targeting refers to testing the range of difficulty of the items. Person-item maps were examined to determine if there were any floor or ceiling effects on the ABLE scale.

Establishing item difficulty allows for several outcomes: the order of difficulty of the items can proceed in an increasing manner; gaps in difficulty level can be detected in the scale; and data is now at an interval-level. The item difficulty also revealed if there were any redundant items on the scale, resulting in the possible reduction of some items. When a Rasch analysis is used, the item difficulty is represented as a logit. The logit is the "natural logarithm of the odds of a person being able to perform a particular task" (Duncan et al., 2003). Once the logit for each item is calculated, the items can then be represented on an item map. This allows the researcher to see any redundant items, as well as gaps in the item difficulty, and the overall hierarchy of the items (Bond & Fox, 2001).

In Rasch analysis, the fit of an item is the estimate of that item's ability to measure a single construct, known as unidimensionality (Bond

& Fox, 2001). For the ABLE scale, the construct being measured is balance. Rasch analysis uses mean square fit statistics to assess whether an item is performing as expected (Bond & Fox, 2001, Elliot et al., 2006). For items with adequate fit, persons with low ability will score high on easy items, and will score low on more difficult items, and persons with high ability will score high on more difficult items. The in-fit mean squares were examined to test for unidimensionality, and the out-fit mean squares were examined when assessing item redundancy (Elliot et al., 2006). Items with an in-fit mean square of >1.4 were considered to test a different construct (Elliot et al., 2006, Bond & Fox, 2001). For the assessment of item redundancy, items with an out-fit value of <0.6 or >1.4 were targeted for further assessment to determine if they should be modified or removed from future versions of the ABLE scale (Bond & Fox, 2001, Elliot et al., 2006).

While targeting and item difficulty focused on the individual items, the Rasch analysis also provided an indication of the instrument's ability to distinguish between individuals with distinct levels of ability which is referred to as the person separation, and tests the spread of the persons on the scale. Person separation (G) is the "ratio of the square root of the variance explained by the measurement model to that of the unexplained variance" (Elliott et al., 2006). The person separation index determines the number of distinct strata that are differentiated by the items. The person separation index was calculated through the formula: strata= [4G +1]/3 (Elliott et al., 2006, Duncan et al., 2003, Bond & Fox, 2001). An alternative method for determining how well the scale can differentiate subjects' abilities along the continuum is to calculate the person separation reliability. This reliability index is based upon the same concept as Cronbach's alpha, which tests for internal consistency, and falls between 0 and 1, with 1 indicating perfect reliability.

Each item on the ABLE scale has distinct definitions for each rating scale category, so that a score of 1 on one item is not equal to a score of 1 on a different item. Therefore, the transition between rating scale categories can differ from one item to the next (Bode, 2001). Thus, the partial credit model was used for the Rasch analysis. This model allows the rater to determine how correct the subject's performance on an item was. When using the partial credit model, in order to correctly place the items on the scale according to level of difficulty, the rating scale categories need to be aligned. Pivot anchoring is a process of aligning these differently-worded rating scale categories to assist in defining the difficulty of each item. Pivot anchoring consists of first assigning a point in each item's rating scale in which the categories represent passing or failing an item. These points are then anchored to a common value for all items on the scale, and then the item difficulties are recalibrated across the scale (Bode, 2001).

All demographic data was analyzed with the Statistical Software for the Social Sciences (SPSS) version 14.0. The Rasch analysis was completed using the WinSteps Software version 3.68.2.

Chapter IV

RESULTS

Demographics:

A total of 104 subjects participated in this study. Each subject was tested once on the ABLE scale for inclusion in the Rasch analysis. Table 1 summarizes the demographic characteristics of the subjects. Subjects were stratified based upon functional ability into three distinct categories. Subjects who were unable to stand or walk (n=42) were classified as "wheelchair users," subjects who could stand for at least 10 seconds with minimal to no physical assistance (n=30) were classified as "standers," and subjects who could ambulate at least 20 feet without an assistive device or physical assistance (n=32) were classified as "walkers."

Table 1. Demographics Results for 104 subjects with SCI.

Characteristic	n= 104
[redacted]	
Gender: n(%)	
Male	79 (76)
Female	25 (24)
[redacted]	
Type of SCI, n (%)	
Motor complete	17 (16)
Motor incomplete	87 (84)
Tetraplegia	59 (57)
Paraplegia	45 (43)
[redacted]	

Targeting and Item Difficulty:

Rasch analysis places item difficulty and person ability along the linear continuum of a logit scale. Figure 1 is a person-item map which displays the item difficulty and person ability of the ABLE scale for 104 subjects with SCI. To the left of the dotted line are the person ability measures (the # symbol represents 2 subjects) and to the right of the dotted line are the item measures, with each item represented by its corresponding number on the ABLE scale. The subjects with the lowest balance ability are located at the bottom of the scale, while the subjects with the highest ability are located at the top of the scale. Similarly, the most difficult items are positioned at the top of the scale, and the easiest items are located at the bottom of the scale.

Figure 1. Person-item map for 28 ABLE scale items as tested on 104 subjects with SCI.

```
              PERSONS - MAP - ITEMS
                   <more>|<rare>
  7                     #  +
  6                    .# T+
  5                 .####  +T
  4                    ##  +  21a     21b     28
  3                  .### S+S 19      23      24      25      26
  2                 #####  +  18      20      27
  1                 #####  +  16      22
  0                 ##### M+M 12      13      14      15      17      9
 -1                   ###  +  10      11
 -2                 #####  +  7       8
 -3                   ###  +S 3a      3b
 -4             ######### S+  1       2       4       5       6
 -5                   .##  +T
 -6                     .  +
 -7                     # T+
 -8                     .  +
                   <less>|<frequ>
 EACH '#' IS 2.
```

Targeting compares the range of item difficulties to the range of person abilities. In Figure 1, the range of person abilities is from 7.95 to -8.41, and the range in item difficulties is from 3.76 to -4.42. Based upon these ranges of ability, a ceiling effect for subjects with abilities greater than 4 logits was noted, as there are no items at these levels to capture their abilities. Similarly, a floor effect was noted for subjects with abilities less than -4 logits, as there are no items at these lesser calibrations.

Item difficulty places the items in order of ascending difficulty, determines if there are any gaps in the levels of difficulty across the scale, and determines if there is any redundancy in the item difficulties. As can be seen in Figure 1, the easiest items are 1, 2, 4, 5, and 6, and are located on the -4 logit. The most difficult items are 21a and b, and 28. Between -4 and 4 logits there are no gaps in item difficulty. However, there are apparent item redundancies at every logit interval, with multiple redundancies at the -4, 0, 2, 3, and 4 logits.

Effect of Pivot Anchoring on Targeting and Item Difficulty:

Pivot anchoring defines a point on each item's rating scale at which a subject would be rated as a pass or fail. For the ABLE scale, passing for each item was defined as the ability to complete the specified task according to the item's instructions, without physical assistance or supervision. Failing was defined as the ability to only partially complete the specified task, or to complete the task with the use of supervision or physical assistance. Using these definitions, pivot points were defined for each item's rating scale (Table 2).

Table 2. Pivot point anchors for each ABLE scale item.

Item	Pivot Point
[redacted]	
2. Seated forward reach	2. able to reach forward 2 inches independently
[redacted]	
4. Pick up/touch object on floor from sitting	3. able to pick up/touch the cup independently but uses arms for support
[redacted]	
6. Posterior seated external perturbations	4. trunk remains steady through all three nudges
[redacted]	
8. Support surface displacement in a wheelchair	2. able to raise basketball to 90 degrees shoulder flexion with elbows extended and maintains or recovers balance during turns in both directions
[redacted]	
10. Static standing balance	4. able to stand ≥1 minute independently
[redacted]	
12. Static standing balance with eyes	4. able to stand ≥ 30 seconds independently

closed	with normal sway
[redacted]	
14. External perturbations in standing	3. utilized hip and ankle strategies to maintain balance, with feet shoulder width apart
[redacted]	
16. Pick up/touch object from floor in standing	4. able to pick up/touch cup independently without use of arms for balance
[redacted]	
18. Turn 180 degrees	2. able to turn 180 degrees in at least 1 direction, in >4 seconds
[redacted]	
20. Balance in tandem/stride stance	4. able to independently achieve and maintain tandem stance ≥30 seconds
[redacted]	
22. Walking over level surface	2. able to walk 20' with an assistive device, independently
[redacted]	
24. Walking with change in direction	3. able to turn direction with minimal hesitation without loss of balance, independently

[redacted]	
26. Walking while carrying an object with 2 hands	3. cadence slows slightly while holding object but ambulates independently
[redacted]	
28. Walking up/down an incline	2. able to walk both up and down ramp slowly, with minimal path deviations, independently

Figure 2 represents the person-item map that resulted from pivot anchoring. As can be seen in this figure, the revised person ability range is from 7.76 to -8.65 logits, and the revised range in item difficulties is 6.05 to -7.06 logits. Based upon these findings, a slight ceiling effect still exists, as there are no items to measure the two subjects with abilities greater than 6 logits. There is also a slight floor effect, as there are no items to measure the one subject with an ability of less than -7 logits.

Figure 2. Person-item map for 28 ABLE scale items as tested on 104 subjects with SCI after pivot anchoring.

```
            PERSONS - MAP - ITEMS
                 <more>|<rare>
  7                  #  +
  6                 .# T+T 21a
  5               .###  +  21b
  4                ###  +  23      28
  3                 ## S+S 19      22      24      26
  2             .#####  +  20      25      27
  1              #####  +  16
  0               ####  +M 11      12      15      18      3a      3b      8
 -1               #### M+  13      17      9
 -2             .#####  +  14
 -3                .##  +S 10      2
 -4          ######## S+   6       7
 -5                ###  +  1
 -6                 ##  +T
 -7                  # T+  4       5
 -8                  .  +
                 <less>|<frequ>
EACH '#' IS 2.
```

As a direct result of the pivot anchoring, adjustments were noted in the item difficulty order. The most difficult item after the pivot anchoring is 21a (right single leg stance), with a difficulty of 6.05 logits. The easiest two items are 4 (pick up/touch an object from the floor from a seated position) and 5 (scooting forward in a chair), both of which are located on the -7 logit. Between the -7 and 6 logits there is one gap in item difficulty, occurring at the -6 logit. Several item redundancies are still noted following the pivot anchoring, with multiple redundancies existing at the -1, 0, 2, and 3 logits.

Examination of Item Fit:

Examining the fit of an item can help identify items which measure a different construct, and can aide in eliminating item redundancy. Table 3 shows the in-fit and out-fit mean square values for all of the items on the ABLE scale. Two items, item 7 (transfers) and item 8 (seated wheelchair perturbations) were determined to have in-fit mean squares of >1.4, suggesting that these items may be measuring a construct other than functional balance.

Table 3. Mean square values for each item on the ABLE scale.

Item	**In-fit**	**Out-fit**
[redacted]		
2	.78	.30
[redacted]		
3b	1.01	2.98
[redacted]		

5	1.07	.79
[redacted]		
7	1.66	1.85
[redacted]		
9	.53	.41
[redacted]		
11	.68	.36
[redacted]		
13	.57	.40
[redacted]		
15	.59	.50
[redacted]		
17	.79	.76
[redacted]		
19	.75	.41

[redacted]		
21a	1.22	.78
[redacted]		
22	.41	.25
[redacted]		
24	.49	.19
[redacted]		
26	.60	.24
[redacted]		
28	.51	.22

Items with an out-fit of <0.6 are considered to be less efficient in measuring the construct. While these items are not a threat to the validity of the scale, they may produce deceptively high reliability estimates. Ten items had outfit values of <0.6, including item 2 (seated forward reach), item 6 (posterior external perturbations in sitting), item 9 (sit to stand), item 11 (stand to sit), item 13 (standing with feet together), item 15 (standing forward reach), item 18 (turn 180 degrees), item 19 (alternate step test), item 21b (left single leg stance), and item 22 (walking over level surface). Items with an outfit of >1.4 are a greater threat to validity, and represent items that are outliers. Four items had an outfit of >1.4, including items 3 a and b (seated lateral reach to the right and left), item 7 (transfers) and item 8 (seated wheelchair perturbations). Therefore, these items should be tested further using a factor analysis with a larger sample size.

Examination of Rating Scale Categories:

To further determine what changes need to be made to the items of the ABLE scale, each item's rating scale categories, or scoring levels, were analyzed using category probability curves and category thresholds. These category curves and thresholds show "the probability of choosing a given rating scale category for every place along the measured variable" (Elliot et al., 2006). The category probability curves are a graphical representation of the probability of each rating scale category for that item being used based upon overall performance on the ABLE. We should expect that subjects with overall lower performance would use the lower rating scale categories, and subjects with higher performance would use the higher rating scale categories. Since we have a full spectrum of recovery represented in our sample, we would expect that for each item we would see 5 prominent curves on the graph, with each curve representing a point at which that particular rating would be more probable than other ratings for a particular ability level. The category threshold tables display the number of times that particular rating category was used for that item, along with the corresponding percentage for the sample.

Appendix D displays the category threshold tables and probability curves for each of the 28 ABLE scale items. As can be seen from the tables and graphs in the appendix, every item on the ABLE scale has at least one underutilized rating scale category.

Suggested Revisions to the ABLE Scale:

Based upon the examination of the item fit statistics, as well as the rating scale category utilization, several changes are suggested for the ABLE scale. Table 4 represents the suggested revisions for the ABLE scale. These revisions will be made as part of a future study.

Table 4. Suggested Revisions to the ABLE Scale.

Item #	Issue(s)	Change to be made
[redacted]		
6	Out-fit < 0.6; redundant with item 7	Rewrite rating scale 1,2,3
[redacted]		
11	Out-fit < 0.6; redundant with items 3, 8, 12, 15, 18	Rewrite rating scale 1,2
[redacted]		
15	Out-fit < 0.6; redundant with items 3, 8, 11, 12, 15, 18	Rewrite rating scale 1
[redacted]		
19	Out-fit < 0.6; redundant with items 22, 24,	Rewrite rating

	26	scale 2
[redacted]		
22	Out-fit < 0.6; redundant with items 19, 24, 26	Rewrite rating scale 1
[redacted]		
7	In-fit >1.4, out-fit >1.4; redundant with item 6	Rewrite rating scale 3
[redacted]		

Person Separation:

Person separation (G) is the "ratio of the square root of the variance explained by the measurement model to that of the unexplained variance" (Elliott et al., 2006). The person separation index determines the number of distinct strata that are differentiated by the items. The person separation index was calculated through the formula: strata= [4G +1]/3 (Elliott et al., 2006, Duncan et al., 2003, Bond & Fox, 2001). The person separation (G) for the ABLE scale, after pivot anchoring, was determined to be 7.67. This resulted in a person separation index of 10.56, meaning that the ABLE scale was able to distinguish 10 distinct strata. The person reliability was also computed using the WinSteps program, and was determined to be .98. This means that the ordering of the person abilities has a 98% chance of being replicated in future studies with a similar population.

Chapter IV

DISCUSSION

The purpose of this study was two-fold. First was the development an all-inclusive clinical instrument to assess balance in the SCI population. This was accomplished by examining the literature to identify the need for a new clinical instrument to assess balance in the SCI population. Once this was established by the lack of reliable and valid outcome measures that can assess balance across the full spectrum of recovery in SCI, a Delphi technique was utilized to develop the scale (Appendix B). The Delphi technique incorporated the feedback provided by 24 clinical experts and 7 advanced experts in SCI rehabilitation and research, to develop a scale with 28 items, that measured balance across the three functional domains of sitting, standing, and walking. A second pilot study, in which 2 experienced physical therapists administered the ABLE scale on 2 subjects with SCI who were at opposite ends of the spectrum of recovery, indicated that the ABLE scale could be easily implemented in the typical physical therapy clinic (Appendix C).

The second purpose of the study was to determine the initial properties of targeting, item difficulty, and person separation of this new scale. This was accomplished through the use of the Rasch analysis. The initial analysis revealed dramatic floor and ceiling effects, thus indicating an inadequate targeting range of the ABLE scale in relation to this population. While this analysis placed items in order of level of difficulty, and no gaps were observed in levels of difficulty between the -4 and 4 logits, there were multiple difficulty levels which exhibited item redundancy.

When using a partial credit model, as was used in the ABLE scale, it is recommended that pivot anchoring be applied to the analysis (Bode, 2001; Elliott et al., 2006). The use of pivot anchoring aligns the rating scale categories' pass points of the items so that they are comparable. This allows for a more accurate analysis of the item difficulties (Bode, 2001). As can be seen through a comparison between Figure 1 and Figure 2, the use of pivot anchoring improved the targeting ability of the ABLE scale, by reducing the floor and ceiling effects. In the initial analysis, there were a total of 14 subjects whose ability levels were greater than the most difficult item, and 7 subjects whose ability levels were less than the easiest item. After the pivot anchoring, there were only 2 subjects whose ability levels were greater than the most difficult item, and 1 subject whose ability level was less than the easiest item.

The pivot anchoring also resulted in an improved spread of item difficulties. The initial analysis revealed multiple item redundancies clustered around the -4, 0, and 3 logit levels. After pivot anchoring, several logits had only one or two items at that difficulty level, although there were still redundancies seen at the -1, 0, 2 and 3 logits.

Due to these multiple redundancies, and the continued presence of the floor and ceiling effects, the items were examined for misfit. Two items had an in-fit mean square value of >1.4 which suggested that these two items measured a construct other than balance. One of the items was item 7 (wheelchair to chair transfers). Transfers may be considered to measure the construct of mobility rather than balance, thus resulting in the high in-fit statistic. This item also had a large outfit statistic, suggesting that several subjects with lower abilities scored higher than predicted on this item, and several subjects with higher abilities scored lower than predicted on this item.

When analyzing the category probability curves for this item in Appendix D, the rating scale category of 3 was the most utilized. This is because this rating scale category allows for subjects to use their arms to assist with accomplishing the transfer. Many subjects with lower ability levels who were unable to sit unsupported, yet had good strength in their arms, could therefore score a 3 on this item. The use of upper extremity support during this task reduces the task to more of a mobility measure, than a true measure of balance. Therefore, it is recommended that the definition of rating scale category 3 for this item be rewritten to more accurately reflect balance, instead of mobility.

The second item with a high in-fit value was item 8 (support surface displacement while seated in a wheelchair). This item had both a large in-fit (3.05) and a large out-fit (9.90). There are several reasons why this item may have had such a large misfit. First, the item was not based on a typical functional activity, such as the other items on the ABLE scale. Thus, it may have been more difficult for the physical therapist raters to administer and correctly score the item. Second, the highest two rating scales require the subject to raise the ball over his/her head. Subjects who could not raise the ball above their head scored a 2 or lower. As many of the subjects tested had tetraplegia (57%), it is possible that they scored lower on this item due to poor upper extremity control, and not because of a lack of balance.

Seated lateral reach, in both the right and left directions (item 3 a and b), also had a high out-fit mean square value. This item was the only item on the test to assess movements in the frontal plane, as all other items assess balance activities in the sagittal or transverse planes of movement. This may have resulted in the high in-fit values, as these frontal plane movements require different balance abilities than the sagittal plane movements. The high out-fit values, in which several subjects with high balance abilities scored unexpectedly low, may be a result of the definitions of the rating scale categories. When the category probability curves in Appendix D are examined, one can see that fewer subjects could reach >10 inches laterally in either direction (43% on 3a, and 42% on 3b) than could reach >10 inches forward (73% on item 2). Thus, it is recommended that the rating scale categories be rewritten for item 3 a and b to more accurately reflect the level of difficulty of this task.

Ten items were found to have a low out-fit mean value (<0.6). While these low out-fit values may result in inflated reliability estimates for the scale, they should not affect the validity of the tool. However, to reduce the redundancy of the items on this scale, these out-fit values may be used to remove misfitting items. Prior to removing the items from the scale, however, the rating scale categories for each of the items should be scrutinized to determine if rewriting the scoring criteria would improve the accuracy of the item. Removing items after only conducting a Rasch analysis may result in the elimination of items which are sensitive to changes in balance over time. Thus, it is recommended that a follow-up study using factor analysis be completed. The use of a factor analysis in a future study will further establish the fit of the items in the ABLE scale by identifying misfitting items, and establish the unidimensionality of the scale.

At first glance, the order of item difficulty was surprising. When developing the scale, the primary author had placed sitting unsupported as item 1, after making the assumption that sitting in a static position was the easiest item on the test. However, the results of the Rasch analysis indicate that items 4 (pick up/touch object from the floor from a seated position) and 5 (scooting forward in a chair) were easier than unsupported short sitting. This is perhaps most likely due to the scoring of the items. In order to obtain the highest score on item 1, the subject must be able to maintain a neutral pelvic position for 2 minutes, independently. Many of the subjects tested lacked the ability to attain and maintain a neutral pelvic position. There were even several subjects who had regained the ability to ambulate without an assistive device, who had not recovered the ability to maintain a neutral pelvic position. Item 4 may have been rated as an easier item, due to the description of rating scale 3, where the subject is allowed to use his/her upper extremity to maintain balance while picking up the cup. As could be seen in the category probability curves in Appendix D, this was one of the most utilized categories for this item. Furthermore, this category had an observed average that was out of order for the rating scale, indicating that a rating scale 3 was actually easier for subjects to attain than rating scale 2. Similarly item 5 may have been rated as an easier item, due to the description of rating scale 3, in which the subject may use his/her upper extremities to move both buttocks forward in the chair. There were several subjects who were unable to maintain unsupported short sitting for 2 minutes, either with or without a neutral pelvic position, who were able to move both buttocks forward in a chair due to the upper extremity support allowed in rating scale category 3. Thus, it is recommended that the rating scale categories of these items be rewritten to improve the accuracy and clarity of the scoring.

The order of item difficulty was also surprising at the top end of the scale. The primary author of the ABLE scale had placed walking up and down stairs (item 27) and up and down an incline (item 28) as the last items on the scale, assuming that they would be the most difficult for individuals with SCI. However, the Rasch analysis has identified standing on one leg as the most difficult item on the scale. This may be due to several factors. On the single leg stance items, in order to achieve a score of 4, subjects could not use upper extremity support, and had to maintain single leg stance with the opposite limb at least 2" off of the ground for at least 20 seconds. To achieve a score of three, the subjects had to maintain this stance for at least 10 seconds. While both ambulating up and down stairs, or up and down an incline require the subject to be in single leg stance, the time that the subject must maintain this single leg stance during these activities is minimal. Many subjects had difficulty maintaining single leg stance for longer than 10 seconds, while others could only maintain single leg stance for this period of time on one lower extremity. This may be a result of impaired lower extremity strength affecting a subject's balance. It was also interesting to observe that standing on the right lower extremity was more difficult than standing on the left lower extremity. This is most likely due to the sample tested, as there may have been more subjects with impaired right lower extremity strength. As manual muscle testing or ASIA exams were not performed during the data collection, we are unable to correlate the results of the single leg stance tests with impairments in strength. This would be interesting to test in future studies.

The person separation *reliability* of .98 was excellent, and suggested that these results were highly reproducible. The person separation *index* of 7.67 was also very high. A separation index of 2.0 is considered to be the minimum acceptable value (Elliott et al., 2006). This high index resulted in the statistical identification of 9 distinct strata in person abilities. Given the large spread of abilities measured in this sample, from -8 to 7 logits, it is not surprising that the ABLE scale was able to distinguish 10 distinct strata. However, the strata identified by the Rasch analysis exist in an abstract statistical model, and may not correlate with the functional abilities observed in the clinical world.

Future studies which correlate the ABLE scale to other clinical balance measures may assist in strengthening these findings.

The Strengths and Weaknesses of the ABLE Scale:

The use of the Rasch analysis has provided insight into the strengths and weaknesses of the ABLE scale. While the ABLE scale clearly has high person separation reliability, there are still several weaknesses to this tool. First, although the use of pivot anchoring more accurately identified the difficulty of the items, there still remains a slight floor and ceiling effect. The floor effect is not as concerning as the ceiling effect, as there was only one subject at the bottom of the scale who did not have an item to appropriately measure his balance. This subject had a complete SCI at the C2 neurological level. An individual with this degree of injury severity would not be expected to be able to maintain his or her balance, as there are no muscles that can be voluntarily activated to assist this individual in sitting. Therefore, there is most likely a small population of individuals with SCI on whom it is not appropriate to assess balance. However, with the ceiling effect, there are two subjects whose positive changes in balance will not be able to be measured, as there are no items that are difficult enough to challenge these subjects' balance abilities. This leads one to question why this ceiling effect has occurred. Is it that the spectrum of recovery after SCI is so great, as evidenced by the large spread in person abilities, that one is unable to capture this complete spectrum with just one outcome measure? Or is it simply that the two subjects at the top of the scale have reached full recovery of balance, and no longer present with balance deficits that compromise their functional independence? Of importance to note, neither the subject that scored a perfect score on the ABLE scale, or the other subject whose total score was 119 out of 120 had experienced a fall in the last year. Thus, leading us to infer that if present, these two subjects' balance deficits did not compromise their functional independence. Future studies on the ABLE scale may consider adding a more difficult item to appropriately challenge the balance of these individuals at the highest end of the spectrum of recovery.

A second weakness of the ABLE scale is the item redundancy noted at several logits. Some of this item redundancy may be due to misfitting items. One possible explanation for these misfitting items is the scoring, or rating scale utility, of each item. If the rating scales for each item are not capable of measuring the degree to which each subject possesses the ability to perform the task, then the difficulty and fit of the item may be calculated incorrectly. A brief analysis of the rating scale utility for each item revealed that there were one or two categories for each item that were underutilized (Appendix D). Possible explanations for this underutilization include: the sample did not have a great enough distribution to capture all of these rating scale categories; five categories are unnecessary to measure each of these balance tasks; the description of each of the underutilized categories was not appropriate to describe the ability level for that task. As can be seen in Figure 2, the person ability range is wide-spread, and presents with a normal distribution. Therefore, the underutilized categories are most likely not the result of sampling error. It is certainly possible that five categories are not necessary to capture balance abilities for each of these tasks in the SCI population. This is similar to the findings by Kornetti et al. (2004) in their analysis of rating scale utility of the Berg Balance Scale in the stroke population. These authors found that each item on the BBS had underutilized categories, and therefore combined the underutilized categories, which resulted in an improved ability to distinguish subjects of varying abilities. However, when considering the combination of underutilized categories on a rating scale, one must determine if these combinations make sense (Bond & Fox, 2001; Elliott et al., 2006; G. P. Williams et al., 2005). One must determine that if by collapsing the categories, if distinct ability levels would be combined (G. P. Williams et al., 2005). This could result in a decreased capability to accurately describe a subject's true level of ability to perform that task. Also one must consider if the combination of categories would result in a substantive change to the pivot point of the item (Bond & Fox, 2001). Furthermore, it is unclear what effect the collapsing of underutilized categories would have on the sensitivity of the ABLE scale to detect changes in balance over time.

It is also possible that the underutilization of several of the rating scale categories was due to the inappropriate description of the corresponding ability level. For example, in item 4, a score of 2 (unable to pick up/touch the cup but comes within 1-2 inches of the cup and keeps balance independently) was only used once. Therefore, the developers of this scale need to consider whether these scoring criteria should be rewritten to create a more clearly defined rating scale category. Perhaps subjects were able to reach within 1-2 inches of the cup, but required supervision to maintain balance. These subjects would have had to be scored a 1 (reaches halfway to cup and needs supervision while trying), since a score of 1 is the only category which allows the subject to have supervision.

Limitations of the Study:

There were several limitations to this study. First, all of the subjects tested on the ABLE scale were tested by raters who were experienced in administering balance assessments to the SCI population. It is unclear how these subjects may have been rated by physical therapists with less experience with balance assessment or the rehabilitation of subjects with SCI. The use of less experienced raters may have resulted in poorer person separation reliability, as well as an increased difficulty in distinguishing between the different rating scale categories for each item. As the purpose of this study was to determine the strengths and weaknesses of the ABLE scale, these experienced raters were specifically chosen so that reliable assessments of the subjects could be made, and would not influence the outcome of the study.

A second limitation of the study was that 50 of the 104 subjects tested were assessed by the primary investigator. As an interrater reliability study was not done to specifically compare the primary investigator's reliability with those of the other raters, there is no way to determine if these assessments are comparable. However, the primary investigator had the same level of experience, and met the same inclusion criteria, as the other raters in the study. Furthermore, the sample tested by the primary investigator was stratified so that the 50 subjects were distributed evenly across the three functional ability groups.

A third limitation of this study was the sample size of 104 subjects. While this sample size has been shown to be appropriate for conducting a Rasch analysis of an outcome measure with 20 items, it precluded the ability to perform a factor analysis (Wang et al., 2005). The factor analysis would have been a useful step in further developing the unidimensionality of the scale, to ensure that all of the items on the ABLE scale measure balance, and not another related construct. However, as the primary purpose of the Rasch analysis in this study was to determine the strengths and weaknesses of the ABLE scale, a factor analysis can be conducted in a future study.

Implications for Future Research:

The present study helped to identify several weaknesses of the ABLE scale. Future studies are needed to specifically address these weaknesses. The first step to further refining the ABLE scale should be to clarify and rewrite the descriptions of underutilized rating scale categories on all of the misfitting items. While many statisticians may argue that collapsing underutilized categories will improve the utility of an item's rating scale, all of them agree that the combination of the rating scale categories must make sense (Bond & Fox, 2001, Elliot et al., 2006, Kornetti et al., 2004). For many of the ABLE scale's items, combining two of the categories would not make sense, as each category represents a distinct functional level, or need for supervision versus independence on the task. Furthermore, collapsing categories for several items on the ABLE scale would result in a 3-level scale, whereas the collapsing of categories on other items would result in a 4-level scale. As a result, the items with only a 3-level rating scale would have a lesser impact on the total score than the items with a 4-level rating scale. Also, collapsing rating scales may result in a decreased sensitivity of the ABLE scale to detect changes in balance over time. Thus, the scale developers should consider rewriting categories to establish more distinct ability levels within each item, instead of collapsing categories.

Another recommendation for a future study is a factor analysis on a larger sample of subjects with SCI, to determine if there are any remaining items that are misfitting, and therefore measure a construct other than balance. Items that are found to be misfit could then be removed from the scale, which may result in further decreased item redundancy. This would help to continue to establish the unidimensionality of the scale, and ascertain that the ABLE scale solely measures the construct of balance.

Once these changes have been made to the ABLE scale, and the item difficulty, rating scale utility and unidimensionality of the scale have been improved; further research should be conducted to examine other psychometric properties. Intra- and interrater reliability should be established for the ABLE scale in the SCI population, using both experienced and novice clinicians. This will ensure that any physical therapist who wishes to use the ABLE scale to assess the balance of a client with SCI may do so reliably, regardless of level of experience. Further studies to establish the concurrent validity of the ABLE scale with other currently utilized outcome measures, including the BBS, the POMA, and the MFRT, should be conducted. Also, fall incidence and performance on the ABLE scale should be correlated to determine if the ABLE scale is accurately able to distinguish and predict fallers in the SCI population.

Chapter VI

CONCLUSION

This study was the first step in the development and testing of a new outcome measure to assess balance in the SCI population. The development of the ABLE scale was completed through the use of a Delphi technique, and the initial testing of the scale was done through a Rasch analysis. The Rasch analysis revealed several initial strengths of the ABLE scale, in particular a high person separation reliability and the ability to distinguish 10 distinct functional strata. The Rasch analysis also revealed several weaknesses of the scale, including slight floor and ceiling effects, one gap in difficulty level across the scale, multiple item redundancies at 4 levels, and several misfitting items.

There continues to be a strong need for a reliable and valid outcome measure to assess balance across the full spectrum of recovery in the SCI population. The steps taken in this study to develop and test the ABLE scale have begun to address this need. Future work on this scale will hope to establish an outcome measure that is the gold standard for balance assessment in the SCI population.

References

Adegoke, B. O. A., Ogwumike, O. O., & Olatemiju, A. (2002). Dynamic balance and level of lesion in spinal cord injured patients. *Afr J Med Med Sci, 31*, 357-360.

Allison, L., & Fuller, K. (2001). Balance and vestibular disorders. In D. A. Umphred (Ed.), *Neurological rehabilitation* (Fourth ed., pp. 616-660). St. Louis, Missouri: Mosby.

Andersson, Å., Kamwendo, K., Seiger Å, & Appelros, P. (2006). How to identify potential fallers in a stroke unit: Validity indexes of four test methods. *Journal of Rehabilitation Medicine, 38*(3), 186-191.

Ardolino, E. M., VanHiel, L., Kern, M., Atkinson, D., Dave, M., Oschner, J., Peschel, C., Lorenz, D., & Behrman, A. L. *Recovery status and compensation effect variability and magnitude of outcome measures with activity-based rehabilitation in individuals with chronic incomplete SCI.* Unpublished manuscript.

Askim, T., Mørkved, S., & Indredavik, B. (2006). Does an extended stroke unit service with early supported discharge have any effect on balance or walking speed? *Journal of Rehabilitation Medicine, 38*(6), 368-374.

Au-Yeung, S., Ng, J., & Lo, S. K. (2003). Does balance or motor impairment of limbs discriminate the ambulatory status of stroke survivors? *American Journal of Physical Medicine & Rehabilitation, 82*(4), 279-283.

Bainbridge, N. J., Davie, M., & Haddaway, M. J. (2006). Bone loss after stroke over 52 weeks at os calcis: Influence of sex, mobility and relation to bone density at other sites. *Age & Ageing, 35*(2), 127-132.

Baloh, R. W., Ying, S. H., & Jacobson, K. M. (2003). A longitudinal study of gait and balance dysfunction in normal older people. *Arch Neurol, 60*(835), 839.

Bastille, J. V., & Body, K. M. (2004). A yoga-based exercise program for people with chronic poststroke hemiparesis. *Physical Therapy, 84*(1), 33-48.

Behrman, A. L., & Harkema, S. J. (2000). Locomotor training after human spinal cord injury: A series of case studies. *Phys Ther, 80*(7), 688-700.

Behrman, A. L., Light, K. E., & Miller, G. M. (2002). Sensitivity of the Tinetti gait assessment for detecting change in individuals with Parkinson's disease. *Clin Rehabil, 16*, 399-405.

Belgen, B., Beninato, M., Sullivan, P. E., & Narielwalla, K. (2006). The association of balance capacity and falls self-efficacy with history of falling in community-dwelling people with chronic stroke. *Arch Phys Med Rehabil, 87*, 554-561.

Benaim, C., Perennou, D. A., Villy, J., Rousseaux, M., & Pelissier, J. Y. (1999). Validation of a standardized assessment of postural control in stroke patients: The postural assessment scale for stroke patients (PASS). *Stroke, 30*, 1862-1868.

Beninato, M., Portney, L. G., & Sullivan, P. E. (2009). Using the international classification of functioning, disability and health as a framework to examine the association between falls and clinical assessment tools in people with stroke. *Phys Ther, 89*(8), 816-825.

Berg, K., Wood-Dauphinee, S., Williams, J. I., & Gayton, D. (1989). Measuring balance in the elderly: Preliminary development of an instrument. *Physiotherapy Canada, 41*(6), 304-311.

Berg, K. O., Maki, B., Williams, J. I., Holliday, P. J., & Wood-Dauphinee, S. L. (1992). Clinical and laboratory measures of postural balance in and elderly Population. *Arch Phys Med Rehabil., 73*, 1073-1080.

Berg, K. O., Wood-Dauphinee, S. L., Williams, J. I., & Maki, B. (1992). Measuring balance in the elderly: Validation of an instrument. *Canadian J Public Health, 83*, S7-11.

Bode, R. K. (2001). Partial credit model and pivot anchoring. *Journal of Applied Measurement, 2*(1), 78-95.

Bogle Thorbahn, L. D., & Newton, R. A. (1996). Use of the berg balance test to predict falls in elderly persons. *Phys Ther, 76*(6), 576-585.

Bond, T. G., & Fox, C. M. (2001). *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahweh, NJ: Lawrence Erlbaum Associates, Publishers.

Brotherton, S. S., Krause, J. S., & Nietert, P. J. (2007). Falls in individuals with incomplete spinal cord injury. *Spinal Cord, 45*, 37-40.

Bruin, E. D., & Murer, K. (2007). Effect of additional functional exercises on balance in elderly people. *Clinical Rehabilitation, 21*(2), 112-121.

Brusse, K. J., Zimdars, S., Zalewski, K. R., & Steffen, T. M. (2005). Testing functional performance in people with parkinson disease. *Phys Ther, 85*(2), 134-141.

Bruyere, O., Wuidart, M., Di Palma, E., Gourlay, M., Ethgen, O., Richy, F., & Reginster, J. (2005). Controlled whole body vibration to decrease fall risk and improve health-related quality of life of nursing home residents. *Arch Phys Med Rehabil. 86*(2), 303-307.

Carr, J. H., Shepherd, R. B., Nordholm, L., & Lunne, D. (1985). Investigation of a new motor assessment scale for stroke patients. *Phys Ther, 65*, 175-180.

Cattaneo, D., De Nuzzo, C., Fascia, T., Macalli, M., Pisoni, I., & Cardini, R. (2002). Risks of falls in subjects with multiple sclerosis. *Arch Phys Med Rehabil, 83*, 864.

Cattaneo, D., Regola, A., & Meotti, M. (2006). Validity of six balance disorders scales in persons with multiple sclerosis. *Disability and Rehabilitation, 28*(12), 789-795.

Chang, W., & Chan, C. (1995). Rasch analysis for outcomes measures: Some methodological considerations. *Arch Phys Med Rehabil, 76*, 943-939.

Charlifue, S., & Gerhart, K. (2004). Community integration in spinal cord injury of long duration. *NeuroRehabilitation, 19*(2), 91-101.

Chern, J., Yang, S., & Wu, C. (2006). Whole-body reaching as a measure of dynamic balance in patients with stroke. *American Journal of Physical Medicine & Rehabilitation, 85*(3), 201-208.

Chiu, A. Y. Y., Au-Yeung, S. S. Y., & Lo, S. K. (2003). A comparison of four functional tests in discriminating fallers from non-fallers in older people. *Disabil and Rehabil, 25*(1), 45-50.

Cho, B., Scarpace, D., & Alexander, N. B. (2004). Tests of stepping as indicators of mobility, balance, and fall risk in balance-impaired older adults. *J Am Geriatr Soc, 52*, 1168-1173.

Chou, C., Chien, C., Hsueh, I., Sheu, C., Wang, C., & Hsieh, C. (2006). Developing a short form of the berg balance scale for people with stroke. *Phys Ther, 86*(2), 195-204.

Cipriany-Dacko, L. M., Innerst, D. A., Johannsen, J., & Rude, V. (1997). Interrater reliability of the Tinetti balance scores in novice and experienced physical therapy clinicians. *Arch Phys Med Rehabil, 78*(1160), 1164.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Conner-Kerr, T., & Templeton, M. S. (2002). Chronic fall risk among aged individuals with type 2 diabetes. *Ostomy Wound Management, 48*(3), 28.

Conradsson, M., Lundin-Olsson, L., Lindelof, N., & et al. (2007). Berg balance scale: Intrarater test-retest reliability among older people dependent in activities of daily living and living in residential care facilities. *Phys Ther, 87*, 1155-1163.

Conradsson, M., Lundin-Olsson, L., Lindelöf, N., Littbrand, H., Malmqvist, L., Gustafson, Y., & Rosendahl, E. (2007). Berg balance scale: Intrarater test-retest reliability among older people dependent in activities of daily living and living in residential care facilities. *Physical Therapy, 87*(9), 1155-1163.

Corriveau, H., Hébert, R., Raîche, M., & Prince, F. (2004). Evaluation of postural stability in the elderly with stroke. *Arch Phys Med Rehabil, 85*(7), 1095-1101.

Daly, J. J., Roenigk, K., Holcomb, J., Rogers, J. M., Butler, K., Gansen, J., McCabe, J., Fredrickson, E., Marsolais, E. B., & Ruff, R. L. (2006). A randomized controlled trial of functional neuromuscular stimulation in chronic stroke subjects. *Stroke (00392499), 37*(1), 172-178.

Datta, S., Lorenz, D. J., Morrison, S., Ardolino, E., & Harkema, S. J. (2009). A multivariate examination of temporal changes in berg balance scale items for patients with ASIA impairment scale C and D spinal cord injuries. *Arch Phys Med Rehabil, 90*(7), 1208-1217.

Desrosiers, J., Rochette, A., & Corriveau, H. (2005). Validation of new lower-extremity motor coordination test. *Arch Phys Med Rehabil, 86*, 993-998.

Di Fabio, R. P., & Seay, R. (1997). Use of the "fast evaluation of mobility, balance, and fear" in elderly community dwellers: Validity and reliability. *Phys Ther, 77*(9), 904-917.

Dijkers, M. (2005). Quality of life of individuals with spinal cord injury: A review of conceptualization, measurement, and research findings... 3rd international congress on the restoration of (wheeled) mobility in SCI rehabilitation: State of the art III, Amsterdam, April 2004. *Journal of Rehabilitation Research & Development, 42*(3), 87-110.

Ditunno, J. F. J., Ditunno, P. L., Graziani, V., Scivoletto, G., Bernardi, M., Castellano, V., Marchetti, M., Barbeau, H., Frankel, H. L., D'Andrea Greve, J. M., Ko, H. Y., Marshall, R., & Nance, P. (2000). Walking index for spinal cord injury (WISCI): An international multicenter validity and reliability study. *Spinal Cord, 38*(4), 234-243.

Dobkin, B. H., Apple, D., Barbeau, H., & et al. (2006). Weight-supported treadmill vs. over-ground training for walking after acute incomplete SCI. *Neurology, 66*(484), 493.

Duarte, E., Marco, E., Muniesa, J. M., Belmonte, R., Diaz, P., Tejero, M., & Escalada, F. (2002). Trunk control test as a functional predictor in stroke patients. *Journal of Rehabilitation Medicine, 34*(6), 267-272.

Duncan, P. W., Bode, R. K., Lai, S. M., Perera, S., & Glycine Antagonist in Neuroprotection Americas Investigators. (2003). Rasch analysis of a new stroke-specific outcome scale: The stroke impact scale. *Arch Phys Med Rehabil, 84*, 950-963.

Duncan, P. W., Weiner, D. K., Chander, J., & Studenski, S. (1990). Functional reach: A new clinical measure of balance. *J Gerontol, 43*, M192-197.

Elliott, R., Fox, C. M., Beltyukova, S. A., Stone, G. E., Gunderson, J., & Zhang, X. (2006). Deconstructing therapy outcome measurement with Rasch analysis of a measure of general clinical distress: The symptom checklist-90-revised. *Psychological Assessment, 18*(4), 359-372.

Eng, J. J., Chu, K. S., Dawson, A. S., Kim, C. M., & Hepburn, K. E. (2002). Functional walk tests in individuals with stroke: Relation to perceived exertion and myocardial exertion. *Stroke (00392499), 33*(3), 756-761.

Eng, J. J., Chu, K. S., Kim, C. M., Dawson, A. S., Carswell, A., & Hepburn, K. E. (2003). A community-based group exercise program for persons with chronic stroke. *Medicine & Science in Sports & Exercise, 35*(8), 1271-1278.

English, C. K., & Hillier, S. L. (2006). The sensitivity of three commonly used outcome measures to detect change amongst patients receiving inpatient rehabilitation following stroke. *Clin Rehabil, 20*, 52-55.

Faber, M. J., Bosscher, R. J., & van Wieringen, P. C. W. (2006). Clinimetric properties of the performance-oriented mobility Assessment. *Phys Ther, 86*(7), 944-954.

Fell, N. T. (2000). Case in point. mental imagery and mental practice for an individual with multiple sclerosis and balance dysfunction... including commentary by Zabolitzki F. *Physical Therapy Case Reports, 3*(1), 3-10.

Forssberg, H., & Nashner, L. M. (1982). Ontogenetic development of postural control in man: Adaptation to altered support and visual conditions during stance. *The Journal of Neuroscience, 2*(5), 545.

Franceschini, M., Di Clemente, B., Rampello, A., Nora, M., & Spizzichino, L. (2003). Longitudinal outcome 6 years after spinal cord injury. *Spinal Cord, 41*(5), 280-285.

Franchignoni, F., Matignoni, E., Ferriero, G., & Pasetti, C. (2005). Balance and fear of falling in Parkinson's disease. *Parkinsonism and Related Disorders., 11*, 427-433.

Franjoine, M. R., Gunther, J. S., & Taylor, M. J. (2003). Pediatric balance scale: A modified version of the berg balance scale for the school-age child with mild to moderate motor impairment. *Pediatric Physical Therapy, 15*(2), 114-128.

Fugl-Meyer, A. R., Jaasko, L., Leyman, I., Olsson, S., & Steglind, S. (1975). The post-stroke hemiplegic patient, I: A method for evaluation of physical performance. *Scand J Rehab Med, 7*, 13-31.

Garland, S. J., Willems, D. A., Ivanova, T. D., & Miller, K. J. (2003). Recovery of standing balance and functional mobility after stroke. *Arch Phys Med Rehabil, 84*(12), 1753-1759.

Gaub, M. G., Prost, E., Bomar, M., Farid, R., Langland, G., & Brown, M. (2004). Efficacy of balance and flexibility intervention in a frail female centenarian. *Journal of Geriatric Physical Therapy, 27*(1), 20-26.

Gavin-Dreschnack, D., Nelson, A., Fitzgerald, S., Harrow, J., Sanchez-Anguiano, A., Ahmed, S., & Powll-Cope, G. (2005). Wheelchair-related falls: Current evidence and directions for improved quality care. *J Nurs Care Qual, 20*(2), 119.

Geiger, R. A., Allen, J. B., O'Keefe, J., & Hicks, R. R. (2001). Balance and mobility following stroke: Effects of physical therapy interventions with and without biofeedback/forceplate training. *Phys Ther, 81*(4), 995-1005.

Grigorenko, A., Bjerkefors, A., Rosdahl, H., Hultling, C., Alm, M., & Thorstensson, A. (2004). Sitting balance and effects of kayak training in paraplegics. *Journal of Rehabilitation Medicine, 36*(3), 110-116.

Hale, L., Schou, E., Piggot, J., Littmann, A., & Tumilty, S. (2003). The effect of a combined exercise programme for people with multiple sclerosis: A case series. *New Zealand Journal of Physiotherapy, 31*(3), 130-138.

Harada, N., Chiu, V., Damron-Rodriguez, J., Fowler, E., Siu, A., & Reuben, D. B. (1995). Screening for balance and mobility impairment in elderly individuals living in residential care facilities. *Phys Ther, 75*(6), 462-469.

Harris, J. E., Eng, J. J., Marigold, D. S., Tokuno, C. D., & Louis, C. L. (2005). Relationship of balance and mobility to fall incidence in people with chronic stroke. *Phys Ther, 85*(2), 150-158.

Hasson, F. (2000). Research guidelines for the Delphi survey technique. *Journal of Advanced Nursing, 32*(4), 1008-1015.

Hatch, J., Body, K. M., & Portney, L. G. (2003). Determinants of balance confidence in community-dwelling elderly people. *Phys Ther, 83*(12), 1072-1079.

Hellström, K., Lindmark, B., & Fugl-meyer, A. (2002). The Falls-efficacy scale, Swedish version: Does it reflect clinically meaningful changes after stroke? *Disabil and Rehabil, 24*(9), 471-481.

Hellström, K., Lindmark, B., Wahlberg, B., & Meyer, A. R. (2003). Self-efficacy in relation to impairments and activities of daily living disability in elderly patients with stroke: A prospective investigation. *Journal of Rehabilitation Medicine, 35*(5), 202-207.

Hsieh, C., Hsueh, I., & Mao, H. (2000). Validity and responsivenss of the Rivermead mobility index in stroke patients. *Scand J Rehab Med, 32*, 140-142.

Hsu, C., & Sandford, B. A. (2007). The Delphi technique: Making sense of consensus. *Practical Assessment, Research and Evaluation, 12*(10), 1-8.

Huang, M., Burgess, R., Weber, M., & Greenwald, N. F. (2006). Performance of balance impaired elders on three balance tests under two visual conditions. *Journal of Geriatric Physical Therapy, 29*(1), 3-7.

Hyndman, D., & Ashburn, A. (2003). People with stroke living in the community: Attention deficits, balance, ADL ability and falls. *Disability & Rehabilitation, 25*(15), 817.

Janssen-Potten, Y. J. M., Seelen, H. A., Drukker, J., Huson, T., & Drost, M. R. (2001). The effect of seat tilting on pelvic position, balance control, and compensatory postural muscle use in paraplegic subjects. *Arch Phys Med Rehabil, 82*(10), 1393-1402.

Janssen-Potten, Y. J. M., Seelen, H. A., Drukker, J., Spaans, F., & Drost, M. R. (2002). The effect of footrests on sitting balance in paraplegic subjects. *Arch Phys Med Rehabil, 83*(5), 642-648.

Janssen-Potten, Y., Seelen, H., Drukker, J., & Reulen, J. (2000). Chair configuration and balance control in persons with spinal cord injury. *Arch Phys Med Rehabil, 81*(4), 401-408.

Juneja, G., Czyrny, J. J., & Linn, R. T. (1998). Admission balance and outcomes of patiens admitted for acute inpatient rehabilitation. *Am J Phys Med Rehabil, 77*, 388-393.

Kamper, D., Parnianpour, M., Barin, K., Adams, T., Linden, M., & Hemami, H. (1999). Postural stability of wheelchair users exposed to sustained, external perturbations. *Journal of Rehabilitation Research & Development, 36*(2), 121-132.

Karatas, M., Çetin, N., Bayramoglu, M., & Dilek, A. (2004). Trunk muscle strength in relation to balance and functional disability in unihemispheric stroke patients. *American Journal of Physical Medicine & Rehabilitation, 83*(2), 81-87.

Kegelmeyer, A. D., Kloos, A. D., Thomas, K. M., & Kostyk, S. K. (2007). Reliability and validity of the Tinetti mobility test for individuals with parkinson disease. *Phys Ther, 87*(10), 1369-1378.

Kizony, R., Raz, L., Katz, N., Weingaden, H., & Weiss, P. L. T. (2005). Video-capture virtual reality system for patients with paraplegic spinal cord injury. *J Rehabil Res Develop, 42*(5), 595-608.

Kornetti, D. L., Fritz, S. L., Chiu, Y., Light, K. E., & Velozo, C. A. (2004). Rating scale analysis of the berg balance scale. *Arch Phys Med Rehabil, 85*, 1128-1135.

Krause, J. S. (2004). Factors associated with risk for subsequent injuries after traumatic spinal cord injury. *Arch Phys Med Rehabil, 85*, 1503-1508.

Landis, J.R., & Koch, G.G. (1977). The measurement of observer *agreement for categorical data. Biometrics*, 33, 159-174.

Langhammer, B., & Stanghelle, J. K. (2003). Bobath or motor relearning programme? A follow-up one and four years post stroke. *Clinical Rehabilitation, 17*(7), 731-734.

Lemay, J., & Nadeau, S. (2009). Standing balance assessment in ASIA D paraplegic and tetraplegic participants: Concurrent validity of the berg balance scale. *Spinal Cord, advance online publication, 22 September 2009* doi:10.1038/sc.2009.119

Leroux, A., Pinet, H., & Nadeau, S. (2006). Task-oriented intervention in chronic stroke. *American Journal of Physical Medicine & Rehabilitation, 85*(10), 820-830.

Li, F., Harmer, P., Fisher, K. J., & McAuley, E. (2004). Tai chi: Improving functional balance and predicting subsequent falls in older persons. *Medicine & Science in Sports & Exercise, 36*(12), 2046-2052.

Liaw, L., Hsieh, C., Lo, S., Chen, H., Lee, S., & Lin, J. (2008). The relative and absolute reliability of two balance performance measures in chronic stroke patients. *Disability & Rehabilitation, 30*(9), 656-661.

Lim, L. I., van Wegen, E. E., de Goede, C. J., & et al. (2005). Measuring gait and gait-related activities in Parkinson's patients own home environments: A reliability, responsiveness and feasibility study. *Parkinsonism Relat Disord, 11*, 19-24.

Lin, M., Hwang, H., Hu, M., Wu, H. I., Wang, Y., & Huang, F. (2004). Psychometric comparisons of the timed up and go, one-leg stand, functional reach, and Tinetti balance measures in community-dwelling older people. *J Am Geriatr Soc, 52*(1343), 1348.

Linacre, J. M. (1999). Investigating rating scale category utility. *Journal of Outcome Measurement, 3*(2), 103-122.

Liston, R. A., & Brouwer, B. J. (1996). Reliability and validity of measure obtained from stroke patients using the balance master. *Arch Phys Med Rehabil, 77*, 425-430.

Lord, S. E., Wade, D. T., & Halligan, P. W. (1998). A comparison of two physiotherapy treatment approaches to improve walking in multiple sclerosis: A pilot randomized controlled study. *Clinical Rehabilitation, 12*(6), 477-486.

Lynch, S. M., Leahy, P., & Barker, S. P. (1998). Reliability of measurements obtained with a modified functional reach test in subjects with spinal cord injury. *Phys Ther, 78*(2), 128-133.

Mackintosh, S., Hill, K., Dodd, K. J., Goldie, P., & Culham, E. (2005). Falls and injury prevention should be part of every stroke rehabilitation plan. *Clinical Rehabilitation, 19*(4), 441-451.

Mahoney, F. I., & Barthel, D. W. (1965). Functional evaluation: The Barthel index. *Md Med J, 14*, 61-65.

Manning, J., Neistadt, M. E., & Parker, S. (1997). The relationship between fear of falling and balance and gait abilities in elderly adults in a subacute rehabilitation facility. *Physical & Occupational Therapy in Geriatrics, 15*(2), 33-47.

Mao, H., Hsueh, I., Tang, P., Sheu, C., & Hsieh, C. (2002). Analysis and comparison of the psychometric properties of three balance measures for stroke patients. *Stroke, 33*, 1022-1027.

Marigold, D. S., Eng, J. J., Dawson, A. S., Inglis, J. T., Harris, J. E., & Gylfadottir, S. (2005). Exercise leads to faster postural reflexes, improved balance and mobility, and fewer falls in older persons with chronic stroke. *Journal of the American Geriatrics Society, 53*(3), 416-423.

McAuley, E., Mihalko, S. L., & Rosengren, K. (1997). Self-efficacy and balance correlates of fear of falling in the elderly. *Journal of Aging & Physical Activity, 5*(4), 329-340.

McKenna, H. P. (1994). The Delphi technique: A worthwhile research approach for nursing? *Journal of Advanced Nursing, 19*, 1221-1225.

Mecagni, C., Smith, J. P., Roberts, K. E., & O'Sullivan, S. B. (2000). Balance and ankle range of motion in community-dwelling women aged 64 to 87 years: A correlational study. *Phys Ther, 80*(10), 1004-1011.

Michael, K. M., Allen, J. K., & Macko, R. F. (2005). Reduced ambulatory activity after stroke: The role of balance, gait, and cardiovascular fitness. *Arch Phys Med Rehabil, 86*(8), 1552-1556.

Middleton, J. W., Sinclair, P. J., Smith, R. M., & Davis, G. M. (1999). Postural control during stance in paraplegia: Effects of medially linked versus unlinked knee-ankle-foot orthoses. *Arch Phys Med Rehabil, 80*(12), 1558-1565.

Mihay, L., Iltzsche, E., Tribby, A., Rushing, K., Spears, J., Wiltfong, H., Schaub, D., & Chronister, A. (2003). Balance and perceived confidence with performance of instrumental activities of daily living: A pilot study of tai chi inspired exercise with elderly retirement-community dwellers. *Physical & Occupational Therapy in Geriatrics, 21*(3), 75-86.

Mihay, L. M., Boggs, K. M., Breck, A. J., Dokken, E. L., & NaThalang, G. C. (2006). The effect of tai chi inspired exercise compared to strength training: A pilot study. *Physical & Occupational Therapy in Geriatrics, 24*(3), 13-26.

Miller, E. W., Quinn, M. E., & Seddon, P. G. (2002). Body weight support treadmill and overground ambulation training for two patients with chronic disability secondary to stroke. *Phys Ther, 82*(1), 53-61.

Monsell, E. M., Furman, J. M., Herdman, S. J., Konrad, H. R., & Shepard, N. T. (1997). Computerized dynamic platform posturography. *Otolaryngology-- Head and Neck Surgery, 117*(4), 394.

Mount, J., Bolton, M., Cesari, M., Guzzardo, K., & Tarsi, J.,r. (2005). Group balance skills class for people with chronic stroke: A case series. *Journal of Neurologic Physical Therapy, 29*(1), 24-33.

Muir, S. W., Berg, K., Chesworth, B., & Speechley, M. (2008). Use of the berg balance scale for predicting multiple falls in community-dwelling elderly people: A prospective study. *Physical Therapy, 88*(4), 449-459.

Musselman, K. E., Fouad, K., Misiaszek, J. E., & Yang, J. F. (2009). Training of walking skills overground and on the treadmill: Case series on individuals with incomplete spinal cord injury... including commentary by Behrman AL with authors' response. *Phys Ther, 89*(6), 601-616.

Nashner, L. M., Black, F. O., & Wall, C. (1982). Adaptation to altered support and visual conditions during stance: Patients with vestibular deficits. *The Journal of Neuroscience, 2*(5), 536.

Nilsagård, Y., Lundholm, C., Denison, E., & Gunnarsson, L. G. (2009). Predicting accidental falls in people with multiple sclerosis -- a longitudinal study. *Clinical Rehabilitation, 23*(3), 259-269.

O'Brien, K., Pickles, B., & Culham, E. (1998). Clinical measures of balance in community-dwelling elderly female fallers and non-fallers. *Physiotherapy Canada, 50*(3), 212.

Olawale, O. A., & Ogunmakin, O. S. (2006). The effect of exercise training on balance in adults patients with post-stroke hemiplegia. *International Journal of Therapy & Rehabilitation, 13*(7), 318-322.

One degree of separation: Paralysis and spinal cord injury in the united states (2009).. http://www.christopherreeve.org/site/c.mtKZKgMWKwG/b.5184189/k.5587/Paralysis_Facts__Figures.htm: The Christopher and Dana Reeve Foundation.

Pal, J., Hale, L. A., & Skinner, M. A. (2005). Investigating the reliability and validity of two balance measures in adults with stroke... including commentary by Tyson S and MacKintosh S. *International Journal of Therapy & Rehabilitation, 12*(7), 308-315.

Paltaama, J., West, H., Sarasoja, T., Wikström, J., & Mälkiä, E. (2005). Reliability of physical functioning measures in ambulatory subjects with MS. *Physiotherapy Research International, 10*(2), 93-109.

Pomeroy, V. M., Evans, B., Falconer, M., Jones, D., Hill, E., & Giakas, G. (2001). An exploration of the effects of weighted garments on balance and gait of stroke patients with residual disability. *Clinical Rehabilitation, 15*(4), 390-397.

Portney, L. G., & Watkins, M. P. (2000). *Foundations of clinical research: Applications to practice* (Second Edition ed.). Upper Saddle River, New Jersey: Prentice Hall Health.

Qutubuddin, A. A., Pegg, P. O., Cifu, D. X., Brown, R., McNamee, S., & Carne, W. (2005). Validating the berg balance scale for patients with Parkinson's disease: A key to rehabilitation evaluation. *Arch Phys Med Rehabil., 86*, 789-792.

Raiche, M., Hebert, R., Prince, F., & Corriveau, H. (2000). Screening older adults at risk of falling with the Tinetti balance scale. *The Lancet, 356*, 1001-1002.

Reuben, D. B., Siu, A. L., & Kimpau, S. (1992). The predictive validity of self-report and performance-based measures of function and health. *J Gerontol, 41*(4), M106-M110.

Robinson, B. S., Gordon, J. M., Wallentine, S. W., & Visio, M. (2002). Effectiveness of physical therapy intervention in decreasing the risk for falls in a community-dwelling aging population. *Orthopaedic Nursing, 21*(1), 55.

Rochester, L., Hetherington, V., Jones, D., Nieuwboer, A., Willems, A., Kwakkel, G., & Van Wegen, E. (2004). Attending to the task: Interference effects of functional tasks on walking in parkinson's disease and the roles of cognition, depression, fatigue, and balance. *Arch Phys Med Rehabil, 85*(10), 1578-1585.

Salbach, N. M., Mayo, N. E., Higgins, J., Ahmed, S., Finch, L. E., & Richards, C. L. (2001). Responsiveness and predictability of gait speed and other disability measures in acute stroke. *Arch Phys Med Rehabil, 82*, 1204-1212.

Sauvage, L.,Jr., Myklebust, B. M., Crow-Pan, J., Novak, S., Millington, P., Hoffman, M. D., Hartz, A. J., & Rudman, D. (1992). A clinical trial of strengthening and aerobic exercise to improve gait and balance in elderly male nursing home residents. *American Journal of Physical Medicine & Rehabilitation, 71*(6), 333-342.

Shirado, O., Kawase, M., Minami, A., & Strax, T. E. (2004). Quantitative evaluation of long sitting in paraplegic patients with spinal cord injury. *Arch Phys Med Rehabil, 85*(8), 1251-1256.

Shore, W. S., deLateur, B. J., Kuhlemeier, K. V., Imteyez, H., Rose, G., & Williams, M. A. (2005). A comparison of gait assessment methods: Tinetti and GAITRite electronic walkway. *J Am Geriatr Soc, 53*(11), 2044-2045.

Shumway-cook, A., & Horak, F. (1986). Assessing the influence of sensory interaction on balance. *Phys Ther, 66*, 1548.

Skulmoski, G. J., Hartman, F. T., & Krahn, J. (2007). The delphi method for graduate research. *Journal of Information Technology Education, 6*, 1-21.

Smedal, T., Lygren, H., Myhr, K., Moe-Nilssen, R., Gjelsvik, B., Gjelsvik, O., & Inger, L. (2006). Balance and gait improved in patients with MS after physiotherapy based on the bobath concept. *Physiotherapy Research International, 11*(2), 104-116.

Smith, J., Forster, A., & Young, J. (2006). Use of the STRATIFY falls risk assessment in patients recovering from acute stroke. *Age and Ageing, 35*, 138.

Smith, P. S., Hembree, J. A., & Thompson, M. E. (2004). Berg balance scale and functional reach: Determining the best clinical tool for individuals post acute stroke. *Clin Rehabil, 18*, 811-818.

Somers, M. F. (2000). *Spinal cord injury: Functional rehabilitation* (2nd ed.). Upper Saddle River, NJ: Prentice Hall.

Soyuer, F., & Öztürk, A. (2007). The effect of spasticity, sense and walking aids in falls of people after chronic stroke. *Disability & Rehabilitation, 29*(9), 679-687.

Spinal cord injury: Facts & figures at a glance. (2008). Birmingham, AL: University of Alabama, National Spinal Cord Injury Statistical Center.

Steffen, T., & Seney, M. (2008). Test-retest reliability and minimal detectable change on balance and ambulation tests, the 36-item short-form health survey, and the unified parkinson disease rating scale in people with parkinsonism. *Phys Ther, 88*(6), 733-746.

Stevenson, T. J. (2001). Detecting change in patients with stroke using the berg balance scale. *Australian J Physiotherapy, 47*, 29-38.

Stevenson, T. J., & Garland, S. J. (1996). Standing balance during internally produced perturbations in subjects with hemiplegia: Validation of the balance scale. *Arch Phys Med Rehabil, 77*, 656-662.

Streveler, R. A., Olds, B. M., & Miller, R. L. (2003). Using a Delphi study to identify the most difficult concepts for students to master in thermal and transport science. Session 2430.

Taylor, L., Whittington, F., Hollingsworth, C., Ball, M., King, S., Patterson, V., Diwan, S., Rosenbloom, C., & Neel, A.,Jr. (2003). Assessing the effectiveness of a walking program on physical function of residents living in an assisted living facility. *Journal of Community Health Nursing, 20*(1), 15-26.

Taylor, L. F., Whittington, F., Hollingsworth, C., Ball, M., King, S. V., Diwan, S., Rosenbloom, C., Patterson, V., & Neel, A. (2003). A comparison of functional outcomes following a physical activity intervention for frail older adults in personal care homes. *Journal of Geriatric Physical Therapy, 26*(1), 7-11.

Teasell, R., McRae, M., Foley, N., & Bhardwaj, A. (2002). The incidence and consequences of falls in stroke patients during inpatient rehabilitation: Factors associated with high risk. *Arch Phys Med Rehabil, 83*, 329-333.

Tinetti M.E., William T.F., Mayewski R. (1986). Fall risk index for elderly patients based on number of chronic disabilities. *Am J Med, 80*(429), 434.

Tinetti, M. E. (1986). Performance-oriented assessment of mobility problems in elderly patients. *J Am Geriatr Soc, 34*(119), 126.

Tinetti, M. E., Baker, D. I., Garrett, P. A., Gottschalk, M., Koch, M. L., & Horwitz, R. I. (1993). Yale ficsit: Risk factor abatement strategy for fall prevention. *J Am Geriatr Soc, 41*(315), 320.

Toole, T., Maitland, C. G., Warren, E., Hubmann, M. F., & Panton, L. (2005). The effects of loading and unloading treadmill walking on balance, gait, fall risk, and daily function in parkinsonism. *NeuroRehabilitation, 20*(4), 307-322.

Trueblood, P. R., Hodson-Chennault, N., McCubbin, A., & Youngclarke, D. (2001). Performance and impairment-based assessments among community dwelling elderly: Sensitivity and specificity. *Issues on Aging, 24*(1), 2-6.

Tsuji, T., Meigen, L., Sonoda, S., Domen, K., & Chino, N. (2000). The stroke impairment assessment set: Its internal consistency and predictive validity. *Arch Phys Med Rehabil, 81*, 863-868.

Tyson, S. F., & DeSouza, L. H. (2004). Reliability and validity of functional balance tests post stroke. *Clin Rehabil, 18*, 916-923.

Urbscheit, N. L., & Wiegand, M. R. (2001). Effect of two exercise programs on balance scores in elderly ambulatory people. *Physical & Occupational Therapy in Geriatrics, 19*(4), 49-58.

van Nes, I. J., Latour, H., Schils, F., Meijer, R., van Kuijk, A., & Geurts, A. C. (2006). Long-term effects of 6-week whole-body vibration on balance recovery and activities of daily living in the postacute phase of stroke: A randomized, controlled trial. *Stroke (00392499), 37*(9), 2331-2335.

Velozo, C. A., & Peterson, E. W. (2001). Developing meaningful fear of falling measures for community dwelling elderly. *Am J Phys Med Rehabil, 80*, 662-673.

Verghese, J., Buschke, H., Viola, L., Katz, M., Hall, C., Kuslansky, G., & Lipton, R. (2002). Validity of divided attention tasks in predicting falls in older individuals: A preliminary study. *J Am Geriatr Soc, 50*(1572), 1576.

Walker, C., Brouwer, B. J., & Culham, E. G. (2000). Use of visual feedback in retraining balance following acute stroke. *Phys Ther, 80*(9), 886-895.

Wang, C., Hsueh, I., Sheu, C., Yao, G., & Hsieh, C. (2004). Psychometric properties of 2 simplified 3-level balance scales used for patients with stroke. *Phys Ther, 84*(5), 430-438.

Wang, W., & Chen, C. (2005). Item parameter recovery, standard error estimates, and fit statistics of the winsteps program for the family of Rasch models. *Educational and Psychological Measurement, 65*(3), 376-404.

Wang, R., Yen, L., Lee, C., Lin, P., Wang, M., & Yang, Y. (2005). Effects of an ankle-foot orthosis on balance performance in patients with hemiparesis of different durations. *Clinical Rehabilitation, 19*(1), 37-44.

Wee, J. Y., Wong, H., & Palepu, A. (2003). Validation of the beg balance scale as a predictor of length of stay and discharge destination in stroke rehabilitation. *Arch Phys Med Rehabil, 84*, 731-735.

Willems, D. A., & Vandervoort, A. A. (1996). Balance as a contributing factor to gait speed in rehabilitation of the elderly. *Physiotherapy Canada, 48*(3), 179-184.

Williams, G., Robertson, V., Greenwood, K., Goldie, P., & Morris, M. E. (2005). The high-level mobility assessment tool (HiMAT) for traumatic brain injury. part 1: Item generation. *Brain Injury, 19*(11), 925-932.

Williams, G. P., Robertson, V., Greenwood, K. M., Goldie, P. A., & Morris, M. E. (2005). The high-level mobility assessment tool (HiMAT) for traumatic brain injury. part 2: Content validity and discriminability. *Brain Injury, 19*(10), 833-843.

Wirz, M., Muller, R., & Bastiaenen, C. (2009 (online)). Falls in persons with spinal cord injury: Validity and reliability of the berg balance scale. *Neurorehabilitation & Neural Repair*, doi:10.1177/1545968309341059

Wolf, B., Feys, H., De Weerdt, W., J, Noom, M., & Aufdemkampe, G. (2001). Effect of a physical therapeutic intervention for balance problems in the elderly: A single-blind, randomized, controlled multicentre trial. *Clinical Rehabilitation, 15*(6), 624-636.

Wood-Dauphinee, S., Berg, K., Bravo, G., & Williams, J. I. (1997). The balance scale: Responsiveness to clinically meaningful changes. *Canadian J Rehabil, 10*(1), 35-50.

Wright, B. D., & Masters, G. N. (1982). *Rating scale analysis*. Chicago: Mesa Press.

Appendix A

The Activity-based Balance Level Evaluation (ABLE Scale)

Purpose: to assess changes in balance across the full spectrum of recovery in the spinal cord injury population.

General Instructions:

- The scale consists of three subscales: sitting, standing, and walking. The scale may be administered in full, or each subscale may be administered and scored separately.
- The participant may be given the option to attempt each task twice. Score the higher of the two attempts.
- The participant may not use an assistive device or bracing for any item on the test, except for items #7 and #22, which allows the participant to use an assistive device only.
- The items should be done in the order listed.
- The examiner must adhere to the instructions provided.
- The examiner must use the equipment as described below.
- If a participant attempts an item, but is unable to perform the activity as per the scoring specifications, the examiner may choose to use the comment box to remark on the participant's performance for future reference.

Equipment:

- 1 standard-height chair without armrests
- 1 standard-sized manual wheelchair with removable armrests
- 1 meter stick/yardstick
- 1 large plastic cup (12-16 oz)
- 1 6-8" step stool
- 1 2x4 block of wood at least 15" long

- 1 inflatable beachball (12" diameter)
- 1 stopwatch
- 1 ADA ramp
- At least 8 standard-height (6-8") steps
- 3 cones or tape to mark walkway

General Definitions:

Safely- the participant performs the task without loss of balance or risk of falling

Loss of Balance- the participant shifts weight out of base of support (BOS) and unable to recover/return to within BOS.

Physical Assistance: The examiner places his/her hands on the participant during an activity in order to provide support, or in some instances, to lift the participant.

> ***Minimal physical assistance***: The examiner places his/her hands on the participant during an activity in order to steady the participant.
>
> ***Moderate physical assistance***: The examiner places his/her hands on the participant in order to prevent the participant from falling, or to help the participant initiate a lift.
>
> ***Maximal physical assistance***: The examiner places his/her hands on the participant in order to lift the participant through the majority of the range of motion.

Supervision: The participant completes the task while the examiner purposefully stands within an arm's reach of the participant, but does not actually touch the participant during the activity.

Independent: The participant safely and successfully completes the task, does not require any physical assistance, and the examiner can stand more than an arm's reach away from the participant.

Demographic and Self-report Items: the purpose of these items is to provide the clinician and researcher with demographic information, as well as to help the examiner determine which subscales may be needed for testing.

A. What is your date of birth?

B. What is your gender?

C. What was the date of your injury?

D. What is the level of your injury?

E. Is your injury complete or incomplete?

- Complete
- Incomplete
- Unsure

F. Do you have sensation below the level of your injury?

- Yes
- No

G. Do you have voluntary movement below the level of your injury?

- Yes
- No

H. Can you feel when you go to the bathroom?

- Yes
- No

I. What percent of your day do you use a wheelchair to get around your home and/or community? Please choose one:

a. I use a wheelchair all of the time, in both my home and community

b. I use a wheelchair sometimes at home, always in my community

c. I use a wheelchair sometimes at home and sometimes in my community

d. I never use a wheelchair at home, and only occasionally in my community (for long distances)

e. I never use a wheelchair at home or in my community

J. Are you able to stand for at least 10 seconds with a little assistance from a caregiver or therapist without bracing and without an assistive device?

- Yes
- No
- Unsure

K. Can you walk 20 feet, with an assistive device if needed, but without bracing and without help from a caregiver?

- Yes
- No
- Unsure

L. How many times have you fallen in the past 12 months (or since your injury, if less than 12 months since injury)? A fall is an event which results in a person coming to rest inadvertently on the ground or other lower level (World Health Organization).

Sitting Balance Subscale:

1. Sitting with back unsupported but feet supported on floor or on a foot stool.

Administration of item: The participant should be seated in a standard height chair without arm rests. The participant should be positioned on the chair so that his/her back is not touching the back of the chair and his/her lower extremities have 90 degrees of flexion in the hips, knees and ankles. If the participant cannot achieve a full neutral pelvis due to an orthopedic condition (ie. Lumbar stenosis, fusion of vertebrae, etc..), have the participant sit as upright as possible, and score appropriately.

Instruction to participant: *Please sit up as straight as you can, with a slight arch in your low back, with your arms folded or resting in your lap for 2 minutes.*

Scoring:

4. able to sit with a neutral pelvis (neither anteriorly nor posteriorly tilted) *independently*, 2 minutes
3. able to sit 2 minutes with posterior pelvic tilt, *independently*
2. able to sit ≥30 seconds with posterior pelvic tilt, with *supervision*
1. only able to sit with posterior pelvic tilt, 10-29 seconds, with *supervision*
0. unable to sit without support ≥ 10 seconds

Comments:

2. **Seated forward reach.**
Administration of item: The participant should be seated on a standard height chair without armrests, leaning against the back of the chair, with his/her sacrum approximately 3" from the back of the chair, so that their back is on an 80° incline. The participant should have 90°of flexion in knees, and ankles, with both feet resting on the floor. A meter stick will be held by another examiner at the height of the participant's shoulder. The participant will flex one shoulder to 90°; the other upper extremity may rest in the participant's lap, but cannot provide support. The ulnar styloid process should be used as a bony landmark for measurement. If the participant is unable to flex either upper extremity to 90°, then both upper extremities can rest in the participant's lap but may not be used for support. In this case, the acromion can be used as the bony landmark for measurement. At no point should the participant touch or rest against the meter stick.

Instruction to participant: *Please raise your preferred arm up to the height of your shoulder. Reach forward as far as possible, and then return to an upright position without using your hands for support. Do not twist your trunk as you reach.*

***Scoring:* Upper extremity used (please circle): R L**

4. able to reach forward >10 inches *independently*
3. able to reach forward 5-10 inches *independently*
2. able to reach forward 2 inches *independently*
1. *able to reach forward but needs supervision*
0. loses balance when trying, requires *physical assistance*

3. Seated lateral reach.
Administration of item: The participant should begin seated in the same position as the seated forward reach test, in a chair without armrests. Prior to reaching laterally, the participant should sit upright, so that his/her trunk is no longer touching the back of the chair. When reaching to the right, the participant should abduct the right shoulder to 90°, and the ulnar styloid process should be used as the bony landmark for measurement. The left upper extremity may rest in the participant's lap, but cannot be used for support. If the participant is unable to abduct the shoulder to 90°, then the acromion may be used as the bony landmark. Repeat with the left upper extremity. Score each upper extremity separately. The patient's hips may come up on the opposite side of the reach.

Instruction to participant: *Please raise one arm up to the height of your shoulder. Reach out to the right as far as possible and return to the middle. Wait 5 seconds then reach out to the left as far as possible and return to the middle. Do not twist your trunk while you reach and keep your feet flat on the floor.*

Right	Left

Scoring: Please mark score in the box provided

4. able to reach >10 inches *independently*

3. able to reach >5 inches with *supervision*
2. able to reach safely 2-5 inches with *supervision*
1. able to attempt but reaches < 2 inches with *supervision*
0. loses balance when trying, requires *physical assistance*

Comments:

4. Pick up/touch an object from the floor from a seated position.
Administration of the item: The participant should begin seated in the same position as the seated forward reach test in a chair without arm rests. A 12-16 oz plastic cup should be placed on the floor, between the participant's feet. Any strategy may be used to pick up the cup, including the use of 2 hands on the cup. If the participant is unable to pick up the cup because of impaired hand function, they may just touch the cup.

Instruction to participant: *Please pick up the cup which is placed in front of your feet, any way you like. Try to use your arms for balance as little as possible.*

Scoring:

4. able to pick up/touch cup *independently* without using arms to maintain balance.
3. able to pick up/touch the cup *independently* but uses arms for support
2. unable to pick up/touch the cup but comes within 1-2 inches of the cup and keeps balance *independently*
1. reaches halfway to cup and needs *supervision* while trying
0. loses balance when trying, requires physical assistance to keep from falling

Comments:

5. Scooting forward in a chair.

Administration of the item: The participant should be seated in a standard-height chair without arm rests with his/her feet in contact with the floor, sitting back as far as possible in the chair so that his/her back is against the backrest. In order to move forward, the participant can either scoot buttocks forward unilaterally or bilaterally. The participant should not push against the back of the chair to slide buttocks forward. The examiner may demonstrate segmentally moving each buttock forward.

Instruction to participant: *Please move your bottom forward to the edge of the chair, using your arms if necessary. Do not push against the back of the chair.*

Scoring:

4. able to move one buttock forward at a time without assistance, without upper extremities
3. able to move both buttocks forward simultaneously with or without upper extremities
2. able to lift buttocks off of chair, but unable to move forward with or without upper extremities
1. requires minimal assistance to lift buttocks and move forward with or without upper extremities
0. requires moderate to maximal assistance to lift buttocks and move forward with or without upper extremities

Comments:

6. Posterior external perturbations in sitting.

Administration of the item: The participant should be seated in a standard-height chair without arm rests with his/her feet in contact with the floor, with arms folded across chest or resting in his/her lap. The participant's sacrum should be ~ 3" from the back of the chair. The participant may not lean against the back of the chair. The examiner gently nudges the participant from the front with one hand on the sternum three times, ensuring that the participant is not dislodged >3". The examiner should apply each nudge 5 seconds apart.

Instruction to participant: *Do not move while I nudge you.*

Scoring:

4. trunk remains steady through all three nudges
3. maintains balance but catches him/herself by placing one or both hands on chair during any of the 3 perturbations
2. maintains balance using any of the above strategies after second push, but falls completely backwards after third push
1. maintains balance using any of the above strategies after first push, but falls completely backwards after second push
0. unable to maintain balance with back unsupported or falls backwards after first push

Comments:

7. Wheelchair to chair transfers.

Administration of item: The participant should be seated in a standard height chair without arm rests. Arrange a standard height/width manual wheelchair with a solid seat and no back cushions (use chair size to keep hip and knee flexion roughly at 90 degrees perpendicular to each other for a stand or squat pivot transfer. The participant may use a sliding board if necessary, but cannot score higher than a 2. The left armrests and footrests may be removed prior by the examiner prior to the transfer.

Instruction to participant: *Please transfer from the chair you are sitting in, to the wheelchair next to you, using your hands as little as possible. Then, when you are ready, please transfer back into the other chair. You may use a sliding board if you need one.*

Scoring:

4. able to *independently* perform a stand pivot/stand step transfer without use of hands
3. able to *independently* perform a stand pivot/stand step transfer with definite need of hands, or performs a squat pivot transfer (participant lifts and laterally scoots bottom by pushing through hands and/or lower extremities) *independently*
2. able to transfer (stand pivot, stand step, squat pivot) with or without a sliding board, with *supervision*
1. needs one person to provide *minimal assist* with or without a sliding board
0. needs one or two people to provide *moderate or maximal assist* with or without a sliding board

Comments:

8. Support surface displacement while seated in a wheelchair.
Administration of the item: The participant should be seated in a standard height/width manual wheelchair as described in item #7. The participant holds a 12 inch diameter inflatable beach ball with both hands and/or wrists, while their feet are supported on wheelchair foot rests. The brake on the left wheel should be locked. Facing the participant, the examiner contacts the top of the propulsion rim on the right side of the wheelchair with the examiner's left hand, while guarding the individual with their right arm. The chair is then turned $1/8^{th}$ of a circle (or 45 degrees) forward in one second by pulling their hand down toward the floor. After a balance response is made or once the participant is returned to upright sitting posture, the examiner returns the propulsion rim rapidly back (45 degrees in one second) to the starting position. The trunk is unsupported during this test and the participant is not allowed to bear weight through their hands on their lap during the test.

Instruction to participant: *Hold the ball with both hands and raise it as high as you can. Keep your trunk still while I turn your chair. Try not to lean against the back of the chair.*

4. able to raise ball over head and maintain or recover balance during turns in both directions
3. able to raise ball over head and maintain or recover balance while turning one direction only
2. able to raise basketball to 90° shoulder flexion with elbows extended and maintains or recovers balance during turns in both directions
1. raises ball ≤6 inches off lap or keeps ball in lap, able to keep trunk still or recover balance during turns in at least one direction
0. unable to sit unsupported for 30 seconds, unable to attempt or tolerate perturbations

Comments:

Standing Subscale: All subjects who are unable to stand would score a zero on items 9 through 28.

9. Arising from a chair.
Administration of item: The participant should begin seated in a standard-height chair without armrests, with back of the knees 6" from the edge of the chair.

Instruction to participant: *Please stand up using your arms as little as possible.*

Scoring:

4. independently arises from chair to full upright standing position without use of arms on first attempt
3. independently arises from chair to full upright standing position with use of arms on first attempt
2. requires two attempts to stand from chair with use of arms
1. able to arise from chair with *minimal assistance*
0. unable or needs moderate to maximal assist to stand

Comments:

10. Static standing balance.
Administration of item: Once in a standing position on a level surface, the participant is instructed to stand with their eyes open without holding on to any devices or people.

Instruction to participant: *Please stand for as long as you can without holding on to anything.*

Scoring:

4. able to stand ≥1 minute *independently*
3. able to stand ≥30 seconds on first attempt with *supervision*
2. able to stand ≥15 seconds on first or second attempt with *supervision*
1. able to stand ≥10 seconds on first or second attempt with *minimal assistance*
0. unable to stand, or stands < 10 seconds with *minimal assistance or greater*

11. Stand to sit.
Administration of item: The participant should transition from a full standing position to a seated position in a standard-height chair without armrests.

Instruction to participant: *Please sit down, try not to use your hands for support.*

Scoring:

4. sits independently, controls descent without use of hands
3. sits independently, controls descent by using legs and/or hands
2. sits independently, but has uncontrolled descent
1. requires minimal assistance to sit safely
0. needs moderate or maximal assistance to sit

Comments:

12. Static standing balance with eyes closed.
Administration of item: The participant should stand on a level surface, with feet hip width apart, without leaning or holding on to any surface with eyes closed.

Instruction to participant: *Please close your eyes and stand still for 30 seconds.*

Comments:

Scoring:

4. able to stand ≥ 30 seconds *independently* with normal sway (uses ankle strategies only)
3. able to stand ≥ 30 seconds safely with minimal excess sway (uses ankle and hip strategies), requires *supervision*
2. able to stand ≥ 10 seconds with moderate excess sway (uses upper extremities to counteract balance), requires *supervision*
1. tolerates eyes closed for <10 seconds but remains standing with *supervision*
0. unable to stand or needs help to keep from falling

13. Static standing balance with feet together and eyes open.
Administration of item: The participant should stand on a level surface without leaning or holding on to any surface and with feet touching so that the medial malleoli of the participant's ankles are in contact with each other. If the participant is unable to place feet completely together due to a biomechanical constraint (such as extreme genu valgum or obesity), then the participant may stand with the medial aspect of the knees touching.

Instruction to participant: *Please move your feet so they are touching each other and stand without holding on to anything.*

Scoring:

4. moves feet together and stands *independently* ≥30 seconds
3. requires *supervision* to move feet together and remain standing for ≥30 seconds
2. needs *minimal assistance* to assume the position but can stand for ≥ 30 seconds, with supervision
1. needs *minimal assistance* to assume the position but can stand for 15 seconds, with supervision
0. unable to stand or requires moderate or maximal assistance to assume or hold the position

Comments:

14. External perturbations in standing.
Administration of item: The examiner gently nudges the participant from the front with one hand on the sternum three times while standing on a level surface with feet shoulder width apart, ensuring that the participant is displaced no more than 3". The examiner should apply each nudge 5 seconds apart. If the participant uses an ankle or hip strategy to independently maintain balance during displacement in this position, then have the participant stand with feet together, as in item #13, and repeat the perturbations.
Instruction to participant: Stand with your feet shoulder width apart (or feet together as indicated). I am going to challenge your balance three times. Try to keep your balance while I nudge you.

Scoring:

4. utilizes ankle and hip strategies to maintain balance with feet together
3. utilizes hip and ankle strategies to maintain balance, with feet shoulder width apart
2. steps backwards and uses legs against chair to maintain balance, with feet shoulder width apart
1. maintains balance after first push but falls into chair after second or third push, with feet shoulder width apart
0. unable to stand or maintain balance/falls into chair after first push, with feet shoulder width apart

Comments:

15. Standing forward reach.

Administration of item: The participant should raise his/her preferred arm to 90°, however he/she should be cued to avoid trunk rotation. The ulnar styloid process is used by the primary examiner as the bony landmark for measurement. If the participant is unable to raise either upper extremity to 90°, then the acromion can be used as the bony landmark for measurement. A ruler should be held by a second examiner at the height of the participant's shoulders on their preferred side. The participant must keep his/her feet still, with heels maintaining contact with the ground while returning to an upright/erect posture, and may not use an assistive device.

Instruction to participant: *Raise your preferred arm to the height of your shoulder. Reach forward as far as you can without falling and without twisting your trunk. Then return to full upright standing. Do not move your feet.*

Comments:

Scoring:

4. able to reach forward ≥ 12 inches *independently*
3. able to reach forward ≥ 6 inches *independently*
2. able to reach forward ≥ 2 inches with *supervision*
1. able to reach forward < 2 inches with *supervision*
0. unable to attempt or requires physical assistance to prevent loss of balance

16. Pick up/touch object from the floor from a standing position

Administration of the item: A 12-16 oz plastic cup should be placed 6" in front of the participant's feet. The participant must begin from a standing position, and must return to a full standing position. Any strategy may be utilized to pick up the cup. If the participant is unable to pick up the cup because of impaired hand function, they may just touch the cup.

Instruction to participant: *Pick up the cup that is in front of your feet any way you like, and stand up with it. Try not to use your hands for support.*

Scoring:

4. able to pick up/touch the cup *independently*, without using arms for balance
3. able to pick up/touch the cup but uses hands for balance and/or requires *supervision*
2. able to bend down to pick up/touch the cup, but requires *minimal assistance* to return to full standing position
1. reaches halfway to cup and needs *supervision* while trying
0. unable to try or loses balance when trying

Comments:

17. Standing trunk rotation.

Administration of item: The participant should stand without leaning or holding on to any surface. A second examiner should stand centered 6" behind the participant's shoulder, opposite to the side of rotation, to encourage a better weight shift. The participant is tested in both directions but only scored once. Please note any cervical or thoraco-lumbar fusion under the comment section.

Instruction to participant: *Turn and look at the other examiner over your left shoulder, while keeping your feet planted. Repeat by looking over your right shoulder.*

Comments:

Scoring:

4. independently rotates shoulders *and* cervical spine each to 90°, in both directions

3. independently rotates shoulders *and* cervical spine each to 90° in one direction only
2. independently rotates shoulders *or* cervical spine separately to <90 °, in both directions
1. requires supervision during rotation
0. requires physical assistance during rotation

18. Turn 180 degrees

Administration of item: The participant may not hold on to anything, and must complete a half circle turn in each direction. The time stops once the participant's feet face exactly opposite to the start position. The participant is tested in both directions but scored only once. The participant can start turning in whatever direction they choose.

Instruction to participant: *While standing, turn around in a half circle, pause for 5 seconds, then turn a half circle back in the other direction.*

Scoring:

4. able to turn 180 degrees independently in ≤4 seconds, in each direction
3. able to turn 180 degrees independently in ≤4 seconds, in one direction only
2. able to turn 180 degrees independently in at least 1 direction, in > 4 seconds
1. needs close supervision or verbal cuing during turning in both directions
0. unable to attempt or needs assistance while turning

Comments:

19. Alternating Step test

Administration of item: Place a 6-8" step/stool 4-6" in front of participant's feet. The participant must alternate placing his/her entire foot on the step while maintaining standing. The participant may not hold on to anything. The examiner counts how many times the participant can place his/her foot on the step in 15 seconds.

Instruction to participant: *Without holding on to anything, alternate tapping each foot on the step/stool as many times as you can in 15 seconds, with the goal of getting 15 foot taps. Do not step up on to the stool.*

Comments:

Scoring:

4. able to complete 15 foot taps in 15 seconds *independently*
3. able to complete 8 foot taps in 15 seconds *independently*
2. able to complete ≥ 4 foot taps in 15 seconds but requires *supervision*
1. able to complete ≥ 2 foot taps in 15 seconds but requires *minimal assistance*
0. unable to attempt, or needs moderate or maximal assistance to keep from falling, or steps with one limb only

20. Balance in tandem/stride stance

Administration of item: The examiner should demonstrate the tandem stance position and alternate stance foot position (step forward with feet shoulder width apart) for the participant. If the participant attempts the tandem stance, and cannot hold the position, he/she may attempt the alternate position. The participant chooses which limb to place forward and is only scored on this one position. The participant is allowed at most 2 attempts to achieve the highest scoring foot position possible, starting each attempt from normal stance position.

Instruction to participant: *Please stand with the heel of one foot directly in front of the toes of the other foot. If you cannot keep your balance in this position, you can take a step forward with one foot, keeping your feet about hip width apart.*

Scoring: **Forward limb (please circle): R L**

4. able to *independently* achieve and maintain tandem stance ≥30 seconds
3. requires *minimal assistance* to achieve tandem stance, but can maintain this position for ≥ 15 seconds with *supervision*

2. able to step forward and maintain stride stance, feet shoulder width apart, ≥30 seconds *independently*
1. requires minimal assistance to step but can maintain this position ≥15 seconds with *supervision*
0. unable to attempt or requires *moderate or maximal assistance* to complete

Comments:

21. Single Leg Stance

Administration of item: The participant must be tested on each leg, and will be scored separately for each leg. The participant may not lean or hold on to any surface during testing.

Instruction to participant: *Stand on your right leg as long as you can without holding on to anything. Please lift your left leg at least 2" off of the ground. Repeat standing on your left leg, lifting your right leg at least 2" off of the ground.*

Right	Left

Scoring: Please mark score in the box provided.

4. able to lift leg at least 2" independently and hold ≥20 seconds
3. able to lift leg at least 2" independently and hold 10 seconds, no contact of weight bearing limb with non-weight bearing limb
2. able to lift leg at least 2" independently and hold 5 seconds, or contacts weight bearing limb with non-weight bearing limb
1. attempts task but is unable to lift ≥2" and/or holds <5 seconds
0. unable to try or needs physical assistance to prevent fall

Comments:

Walking Subscale: For items 22-26, the examination should take place on the same 20 foot level walkway surface consisting of tile or low pile carpeting. The walkway should be cleared of all obstacles. The participant is not allowed physical assistance or use of bracing during these tasks but may use their assistive device on item 22 only . The start and finish of the walkway should be clearly marked with tape or cones.

22. Walking over level surface

Administration of item: The participant should walk 20' over a level surface. The participant may NOT receive physical assistance from the examiner. The participant may use an assistive device as necessary, but cannot score higher than a 2. No bracing is allowed during testing.

Instruction to participant: *Walk at your normal speed from here to the end of the walkway.*

Scoring:

4. able to walk 20' without an assistive device; *independently*, no loss of balance
3. able to walk 20' without an assistive device; with *supervision*, regains balance easily using abducted arms
2. able to walk 20' with an *assistive device, independently*
1. able to walk 20' with an *assistive device and supervision*
0. unable to walk 20' with an assistive device without physical assistance

Comments:

23. Walking with horizontal head turns

Administration of item: The participant should ambulate on the same walkway as the previous item. The participant is asked to turn their head 90 degrees (or to the point of cervical range restriction), maintaining each head position for 3 steps. The examiner is encouraged to demonstrate this item. The participant may not use an assistive device or physical assistance from the examiner.

Instruction to participant: *Begin walking at your normal pace. When I tell you to "look right," keep walking straight, but turn your head to the right . Keep looking to the right until I tell you, "look straight,"* *then keep walking straight, but return your head to the center. When I tell you to "look left," keep walking straight, but turn your head to the left. Keep your head to the left until I tell you "look straight," then keep walking straight, but return your head to the center.*

Scoring:

4. able to maintain constant gait speed while turning head in both directions, *independently*
3. *hesitates slightly* while turning head, but does not lose balance or deviate inside a 15 inch wide path
2. *hesitates considerably* and/or *laterally deviates* within a 15" wide path with head turns, requires supervision
1. *laterally deviates outside* a 15" wide path while turning head, requires supervision
0. unable to try/requires physical assistance to prevent a fall

Comments:

24. Walking with change in direction.

Administration of item: The participant should ambulate on the same walkway as the previous item. The participant may not use an assistive device or physical assistance from the examiner. Place a cone halfway down the walkway. The examiner may demonstrate a smooth turn around the cone. The participant may turn around the cone in either direction. Document the direction of turn in the comment box for reference for future testing.

Instruction to participant: *Please walk to the cone, turn around it without hesitation, and return back to the starting position.*

Comments:

Scoring:

4. able to turn direction *without hesitation* and without loss of balance, *independently*
3. able to turn direction with *minimal hesitation* without loss of balance, *independently*
2. approaches cone, *stops, slowly turns* around cone, without loss of balance, requires *supervision*

1. approaches cone, *stops, loses balance* when turning but does not need physical assistance to prevent fall
0. unable to try/requires physical assistance to prevent fall

25. Stepping over object while walking

Administration of item: The participant should ambulate on the same walkway as the previous item. The participant may not use an assistive device or physical assistance from the examiner. Place a 2x4 piece of wood halfway down the walkway, perpendicular to the walkway. The examiner may demonstrate stepping over the 2x4.

Instruction to participant: *Begin walking at your normal speed. When you come to the piece of wood, please step over it, not around it or on it.*

Comments:

Scoring:

4. able to maintain constant speed while stepping over 2x4.
3. stops, steps over 2x4, does not lose balance
2. able to clear 2x4, loses balance but does not need physical assistance to recover
1. stops, unable to clear 2x4, but does not lose balance
0. unable to try/requires physical assistance to prevent falling

26. Walking while carrying an object with 2 hands

Administration of item: The participant should ambulate the full length of the walkway used for the previous items. The participant may not use an assistive device or physical assistance from the examiner. The object should be a 12" inflatable beachball, or an object of similar size and weight. The participant must carry the ball with 2 hands (clenched fists is acceptable for participants with impaired hand function).

Instruction to participant: *Walk down the walkway at your normal pace while holding this ball with both of your hands.*

Scoring:

Comments:

4. maintains consistent speed while holding object, ambulates *independently*,
3. cadence slows slightly while holding object but ambulates *independently*
2. laterally deviates within a 15" wide path while holding object, requires *supervision*
1. laterally deviates outside a 15" wide path while holding object, or drops object >2 times during one pass, requires *supervision*
0. unable to try/needs physical assistance or an assistive device to a prevent fall

27. Walking up/down stairs

Administration of item: At least 8 standard-height (6-8") steps should be used. The participant may not use an assistive device to complete the task. If more than 10 steps are used, please note the total number of steps that the participant was able to negotiate.

Instruction to participant: *Walk up the stairs with your typical pattern using the rails if you need to for safety. At the top of the stairs, turn around and walk down.*

Scoring:

4. able to walk up and down steps without rail, with reciprocal pattern, *independently*
3. able to walk up and/or down steps with rail with reciprocal pattern, with *supervision*, OR able to walk up/down stairs without rail, with step-to pattern, *independently*
2. *able to walk up and down steps with or without rail, with step-to* pattern, with *supervision*
1. able to walk up and down steps with rail, with step-to pattern, with *minimal physical assistance* in each direction
0. unable to try/requires *moderate or maximal physical assistance*

Comments:

Total # steps: _______

28. Walking up/down an incline
Administration of item: An ADA graded ramp (1 foot of length for every 1 inch of rise), such as an entrance ramp into a building, should be used. The participant is not allowed to use an assistive device.

Instruction to participant: *Please walk up and down the ramp without holding on.*

Scoring:

4. able to walk both up and down ramp, *independently*, at or close to, normal walking speed
3. able to walk both up and down ramp, *independently*, but one direction is at a slower speed
2. able to walk both up and down ramp slowly, with minimal path deviations, *independently*
1. able to walk both up and down ramp slowly, with large path deviations (outside 15" wide path) and requires *supervision*
0. . unable to try/requires an assistive device and/or physical assistance to walk up/down ramp

Comments:

Participant ID: __________

Date: _______________

Rater:_______________

Score Sheets:

Demographic Items:

Date of birth:	
Gender:	
Date of injury:	
Level of injury:	
Completeness of injury:	
Sensation below injury:	
Voluntary movement below injury:	
Sacral sparing:	
Percent of day in a wheelchair:	
Stand for 10 seconds:	
Walk 20 feet:	
Number of falls:	

Sitting Balance Subscale		
Item	Score	Comments
1. Sitting		
2. Seated Forward Reach		
3.a. Seated Lateral Reach (Right)		
3.b. Seated Lateral Reach (Left)		
4. Pick up object in sitting		
5. Scooting forward in chair		
6. Seated external perturbations		
7. Transfers		
8. Wheelchair perturbations		
Sitting Balance Score (___/36)		

Standing Balance Subscale		
Item	Score	Comments
9. Sit to Stand		
10. Standing		
11. Stand to Sit		
12. Standing with eyes closed		
13. Standing with feet together		
14. External perturbations in standing		
15. Standing forward reach		
16. Pick up object from standing		
17. Look over shoulder in standing		
18. Turn 180 degrees		
19. Alternate step-ups		
20. Tandem stance		
21.a. Standing on one leg (right)		
21.b. Standing on one		

leg (left)		
Standing Balance Score (___/56)		

Walking Balance Subscale		
Item	Score	Comments
22. Walking over level surface		
23. Walking with head turns		
24. Walking with change direction		
25. Stepping over object while walking		
26. Walking with object in 2 hands		
27. Walking up/down incline		
28. Walking up/down stairs		
Walking Balance Score (___/28)		
Total Score		

Appendix B

PILOT STUDY 1: THE DELPHI TECHNIQUE

Purpose of the Study:

The means by which researchers have traditionally measured balance in the SCI population, mainly with the use of force plates and EMG data, are unavailable in the typical physical therapy clinic. There are currently no outcome measures that have been developed and validated to specifically assess balance in the SCI population. Therefore, the purpose of the pilot study was to generate items for a new clinical outcome measure, the ABLE scale, using expert consensus.

Subjects:

There were a total of two rounds of the Delphi technique plus a round of advanced critique. Subjects in all three rounds were physical therapists who had at least 5 years of physical therapist practice, at least 2 years of evaluating and treating patients with SCI, and at least 2 years of administering the BBS or the POMA. Subjects in rounds 1 and 2 were recruited from the 14 Model SCI Systems, located nationally in the United States, as well as from the 7 centers of the NeuroRecovery Network (NRN), and the NeuroPT listserve, an electronic mailing list operated by the Neurology Section of the American Physical Therapy Association (APTA). Subjects who participated in the advanced critique were recruited from the Balance Committee of the NRN.

In order to anonymously recruit experts in SCI rehabilitation, a letter was electronically sent to all of the supervisors at each of the 14 Model SCI Systems centers and the 7 NRN centers, asking each supervisor to identify at least 2 physical therapists who met the inclusion criteria for an expert in SCI rehabilitation. The supervisors were asked to forward a letter to each expert, which requested the expert's participation in the study, as well as explained the purpose of the study, and the instructions for taking the online survey. To recruit experts from the NeuroPT listserve, an email containing the same information as in the expert letter, was sent to all members of the listserve. The letter clearly stated the purpose of the study, as well as the inclusion criteria.

Procedure:

Round 1:

Approval for the study was obtained from the Institutional Review Board (IRB) of Seton Hall University. The current version of the ABLE scale, which was written by the primary investigator and reviewed by the NRN Balance Committee, was posted online via Seton Hall University's ASSET survey program. At this time, the ABLE scale consisted of 30 items, which tested balance in the domains of sitting, standing, and walking. All experts recruited for the study were given instructions on how to access the survey via ASSET, and were given 2 weeks to complete the survey. In the survey, the experts were presented with each item of the ABLE scale, and were asked several questions regarding the item, including the importance of including the item in the scale, the clarity of the wording, the appropriateness of the scoring, and the feasibility of administering the item in a physical therapy clinic. Experts were also provided with the opportunity to offer suggestions on improving each item and the scale as a whole.

Round 2:

The results from the first round of the survey were reviewed by the research team. Using an 80 % agreement requirement for an item to be modified or deleted, the ABLE scale was revised. The revised scale, noting the items modified or deleted, was posted online via ASSET. Experts were contacted again, through either the supervisors at the Model SCI Systems centers and the NRN centers, or through the NeuroPT listserve. Experts were instructed to take the second-round survey only if they had completed the first-round survey and were given instructions on how to access the survey. The second survey presented each item of the ABLE scale, and the experts were asked to answer the questions following any item which had been modified.

Advanced Review:

Once the ABLE scale went through a 2-round Delphi review process, a final review was conducted by the NRN Balance Committee. The NRN Balance Committee consists of 7 members, who have not only evaluated and treated individuals with SCI, but have also conducted research on animals and humans with SCI, taught classes on SCI rehabilitation, or have published papers or book chapters on SCI rehabilitation. Thus, these individuals represent a higher level of expertise than the experts surveyed in rounds 1 and 2.

The members were then asked to offer any feedback on the current version of the scale, and the scale was modified based upon these suggestions. The Balance Committee members were then presented with a final version of the ABLE scale. This version was also posted online via ASSET, and the members of the Balance Committee were asked to complete a survey to answer questions regarding the importance of including the item in the scale, the clarity of the wording, the appropriateness of the scoring, and the feasibility of administering the item in a physical therapy clinic. Members could choose not to complete the survey if they were satisfied with the final version of the ABLE scale. The chairperson of the Balance Committee, who is the primary investigator of the study, recused herself from completing the survey.

Results:

Round 1:

Twenty-four experts completed the first round survey. The demographic and expertise information is presented in Table 5. Of the 24 participants, 87.5% were female, 58.3% had more than 10 years of experience as a physical therapist, and 41.7% had 6 to 10 years of experience evaluating and treating individuals with SCI. Participants appeared to have more experience administering the BBS than the POMA, as 37.5% had more than 10 years of experience administering the BBS, whereas only 20.8% had more than 10 years of experience administering the POMA.

The percent agreement necessary to reach a consensus was set at 80%. Of the 30 items presented to the experts, 19 items reached an agreement of at least 80%. Eight of the remaining items were modified based upon the suggestions of the experts. Three items, seated rotational reach, sit to supine, and walking with an object in one hand, were removed from the scale, as the majority of the experts indicated that these items were not important to include in this outcome measure. Thus, at the end of round 1, the ABLE scale consisted of 27 items across three domains.

Table 5. Expert Demographics and Experience- Round 1 (n=24)

	Practicing PT	**Evaluate and Treat SCI**	**Administer BBS**	**Administer POMA**
[redacted]				
2-5 years	0.0%	29.2%	29.2%	33.3%
[redacted]				

More than 10 years	58.3%	29.2%	37.5%	20.8%

Round 2:

In round 2, 21 of the 24 experts completed the survey, resulting in an attrition rate of 12.5%. All of the 8 modified items from round 1 reached a consensus in round 2.

Advanced Review:

All members of the NRN Balance Committee offered feedback on the scale prior to administering the final survey. During this time, one item was added to the sitting balance subscale, and minor editorial changes were made to several other existing items, for improved clarity with scoring. Four members of the NRN Balance Committee completed the online survey. All items on the final version of the ABLE scale reached a consensus of at least 80% agreement. After this round of advanced review, the ABLE scale consisted of 28 items across three domains (Appendix A).

Discussion:

The purpose of using the Delphi technique was to establish the content validity of the ABLE scale through expert consensus. The Delphi allowed for experts across a wide geographical area to be surveyed electronically and anonymously. However, there are limitations to the Delphi technique. It is often criticized for not allowing participants to fully explain their opinions (Hasson, Keeney & McKenna, 2000). Furthermore, one must keep in mind that just because a consensus has been reached, it does not mean that this consensus is the correct answer (Hasson, Keeney & McKenna, 2000). However, we sought to prevent these limitations through the methodology of this study.

This first limitation of a Delphi technique, not allowing participants to fully explain their answers, was prevented through the design of the survey. Participants were given ample space to voice their opinions about the importance, wording, and scoring of each item, as well as to comment on the scale as a whole. The feedback provided by the participants in round 1 was taken into account during the modifications of the scale, and was presented to the participants in round 2.

To prevent the second limitation of the Delphi technique, once the consensus was reached on all items at the end of round 2, the ABLE scale went through a round of Advanced Review. The use of experts in the field of SCI research allowed for a more critical appraisal of the scale. This ensured that the ABLE scale is appropriate not only for clinical use, but for use in the research setting as well.

While content validity was reached after the three rounds of review of the ABLE scale the ABLE scale may be time consuming to administer in a clinical setting as it consists of 28 items measured across the three domains of sitting, standing, and walking. Therefore, further testing is currently under way to refine the ABLE scale. This scale is being tested for redundant items, and to determine if there are gaps in the level of difficulty across the scale, which would make it necessary to either remove, modify, or add new items. In future studies, additional psychometric testing will also be conducted on the scale to determine the reliability, construct and concurrent validity, and minimal detectable change score.

Conclusion:

This study was the first step in the development of the ABLE scale, a new outcome measure to assess balance in the SCI population. This scale is being developed because there are currently no outcome measures that have been proven to be reliable and valid in measuring balance in the SCI population. This study utilized a Delphi technique to generate the testing items and establish the content validity of the new ABLE scale. While the scale has the potential to measure balance abilities across a wide spectrum of recovery after SCI, further testing is required to refine the scale and establish the psychometric properties before it can be used in the clinical or research settings.

Appendix C

Pilot Study 2: The Feasibility of Administering the ABLE Scale in the Incomplete SCI Population

Purpose of the Study:

The Activity-based Balance Level Evaluation (the ABLE scale) was developed through the use of a Delphi technique. This technique resulted in an outcome measure with a total of 28 items measured across the three functional domains of sitting balance, standing balance, and walking balance. While the consensus of experts in the field of SCI rehabilitation and research helped to establish the content validity of the scale, further testing is needed to determine if the ABLE scale can be easily implemented in the typical physical therapy clinic. Thus, the purpose of this study is to test the clarity and feasibility of administering the ABLE scale to individuals with SCI in the typical physical therapy clinic.

Subjects:

Patient Subjects:

Two individuals with incomplete SCI who were currently participating in the Locomotor Training program at Magee Rehabilitation Hospital were recruited for this study. Demographic information on these subjects is presented in Table 6. Subjects were specifically chosen based upon their functional levels. One subject, Subject X, was able to stand and walk without an assistive device. The other subject, Subject Y, was unable to stand without assistance. This disparity in functional level was chosen to allow for a greater discrepancy in the testing of the ABLE scale items.

Table 6. The Demographic Characteristics of Two Subjects with Incomplete SCI.

	Subject X	**Subject Y**
[redacted]		
Gender	Male	Male
[redacted]		
Neurologic Level of Injury	T10	C5

[redacted]

Therapist Raters:

Two physical therapists were recruited to administer and score the individuals with SCI in this study. Both physical therapists worked at Magee Rehabilitation Hospital, and had a minimum of 5 years of physical therapist experience in the evaluation and treatment of patients with SCI. These physical therapists also had at least 2 years of experience in the administration of balance outcome measures in this patient population. Furthermore, both of these physical therapists taught SCI rehabilitation in physical therapy programs in local universities. Both physical therapists consented to participate in this study.

Procedure:

Approval for the study was obtained from the Magee Rehabilitation Hospital Institutional Review Board (IRB). Each patient subject signed an informed consent form. Each patient subject was tested by each therapist rater, on separate days, prior to his scheduled physical therapy time for that day. In each case, testing took place one day apart, to ensure for adequate rest between testing sessions.

Prior to initiating testing, the physical therapist raters were given copies of the ABLE scale, and the procedure for administering and scoring the items was briefly reviewed with the primary investigator. During the testing of each patient subject, the therapist raters were asked to not only score the subject, but also comment on the clarity and ease of administering and scoring each item, as well as the overall flow of the test. The physical therapists also recorded the time it took to administer the entire ABLE scale to each subject. The primary investigator was not present during the testing of any of the subjects, so as not to influence the results.

Data Analysis:

All data was analyzed using SPSS version 14.0. Descriptive statistics were used for demographic data. The interrater reliability was calculated for both subjects together, as well as separately, using Cohen's kappa coefficient. For all estimations of reliability, the kappa values were classified as: .41-.60= moderate agreement, .61-.80= substantial agreement, and .81-1.00= almost perfect agreement (Landis and Koch, 1977).

Results:

Table 7 displays a synopsis of the comments provided by the physical therapist raters on the feasibility of administering the ABLE scale in the clinic, including the time to complete the test on each subject.

Table 7. The Feasibility of Administering the ABLE Scale in a Physical Therapy Clinic to Two Subjects with Incomplete SCI.

Topic	Feedback
[redacted]	
Readability of items	3 items had suggestions for changes in instructions 1 item had question regarding scoring criteria
[redacted]	
Time to complete	Patient X= 50 minutes Patient Y= 10 minutes

As can be seen in Table 8, the interrater reliability ranged from .572 to .631, depending on the subject tested. Subject Y, who was unable to complete the standing and walking items, had a higher reliability than Subject X, who was able to attempt all of the items on the ABLE scale.

Table 8. The Interrater Reliability of the ABLE Scale on Two Subjects with Incomplete SCI as Administered by Two Physical Therapist Raters.

Patient	Cohen's kappa
[redacted]	
Patient X	.572 (p<.0001)
[redacted]	

Discussion:

The purpose of this pilot study was to examine the feasibility of administering the ABLE scale in a typical physical therapy clinic. The results could then be used to make further adjustments to the scale, prior to initiating testing of the scale on a larger sample of individuals with SCI. A secondary purpose of the study was to briefly examine the interrater reliability of the scale, to determine if there were any issues in the replicability of items in this sample population.

Overall, the comments from the physical therapist raters suggested that the ABLE scale could be easily implemented in the SCI population. Both physical therapists felt that item 8 (support surface displacements while seated in a wheelchair) was difficult to administer, due to the length of the instructions, and increased difficulty with understanding how to complete the task. Suggestions were made by the therapist raters to clarify the instructions for this item. The therapists also suggested that the test administrator be instructed to demonstrate the task for item 5 (scooting forward in a chair), as they felt that subjects automatically moved their buttocks forward simultaneously, even if the individual could move one buttock forward at a time. Furthermore, the therapist raters suggested that the instruction to the participant for item 25 (stepping over object while walking) include the phrase "do not step on the block" to ensure that the subject would step over the piece of wood.

Overall, the therapist raters felt that the time to complete the full ABLE scale was somewhat lengthy (50 minutes). The therapists felt that the increased time to complete the scale was due to the time needed to collect and set-up the equipment required, as well as time needed to carefully read the instructions for each item on the scale. Therefore, both physical therapists stated that they felt they could reduce this amount of time in the future, once they had more experience administering the scale.

The interrater reliability for the ABLE scale as calculated for both patient subjects represented substantial agreement. The interrater reliability for Subject X was lower than for Subject Y. This is most likely due to the fact that Subject X was able to attempt all of the items, whereas Subject Y was only able to attempt the sitting balance items. Therefore, there may be more variability in the scoring of the standing and walking balance items. However, the overall interrater reliability of the ABLE scale was substantial (kappa= .627 $p<.0001$), indicating good initial replicability of the administration and scoring of the items on this scale.

There were several limitations to this study. First, the physical therapist raters used were experienced clinicians with a high level of skill in the evaluation and treatment of individuals with SCI. It is unclear how the reliability of the ABLE scale would be affected through the administration and scoring of the scale by novice clinicians.

Another limitation was the small sample size of the study. Both subjects tested were males with chronic, incomplete SCI. Also, each subject was at the extreme end of the spectrum of recovery. However, as the purpose of the study was not to establish the psychometric properties of the scale, but instead to test the feasibility of administering the scale in the clinic, this sample is considered sufficient. Testing on this sample showed that the ABLE scale could be easily administered in a physical therapy clinic to subjects with wide ranges of functional abilities.

Conclusion:

This study tested the feasibility of administering the ABLE scale in the SCI population, in a typical physical therapy clinic. Although this study utilized a small sample and only two physical therapist raters, the results showed that the ABLE scale could be easily and reliably implemented in a physical therapy clinic.

Appendix D

Category Probability Curves

Item 1: Unsupported sitting

ITEM DIFFICULTY MEASURE OF -5.91 ADDED TO MEASURES

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
0	0	7	7	-6.48	-6.83	1.46	1.20	NONE	(-7.40)	0
1	1	5	5	-6.31	-5.72	.66	.94	.00	-6.19	1
2	2	6	6	-5.37	-5.01	1.19	.79	.38	-5.37	2
3	3	25	25	-2.67	-3.63	.95	2.49	.00A	-3.85	3
4	4	59	58	1.61	1.97	2.43	1.90	3.75	(-1.04)	4

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	ZONE		50% CUM. PROBABLTY	COHERENCE M->C	C->M	ESTIM DISCR	
0	NONE		(-7.40)	-INF	-6.80		50%	28%		0
1	-5.91	.58	-6.19	-6.80	-5.77	-6.42	25%	20%	.14	1
2	-5.53	.48	-5.37	-5.77	-4.86	-5.73	25%	66%	1.26	2
3	-5.91	.41	-3.85	-4.86	-2.09	-5.22	55%	44%	1.29	3
4	-2.16	.39	(-1.04)	-2.09	+INF	-2.13	84%	83%	-.06	4

M->C = Does Measure imply Category?
C->M = Does Category imply Measure?

```
        CATEGORY PROBABILITIES: MODES - Structure measures at intersections
P      -+-------+-------+-------+-------+-------+-------+-------+-------+-
R  1.0 +000000000000000000                                        44444+
O      |                 0000                                   4444   |
B   .8 +                     00                              44        +
A      |                       0          333333          44           |
B   .6 +                        0       3        33 44                 +
I   .5 +                         0    3            *                   +
L   .4 +                          0  3           4 33                  +
I      |                         111**         44     3                |
T   .2 +                      11  2***222  44         333              +
Y      |                 11111  2*3  0*1***22            3333          |
    .0 +*************************444444*0*************************+
O      -+-------+-------+-------+-------+-------+-------+-------+-------+-
F      -8      -6      -4      -2       0       2       4       6       8
        PERSON [MINUS] ITEM MEASURE
```

Item 2: Seated forward reach

TEM DIFFICULTY MEASURE OF -4.25 ADDED TO MEASURES

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
0	0	3	3	-6.55	-7.08	1.80	1.27	NONE	(-9.04)	0
1	1	13	13	-6.14	-5.85	.50	.51	-3.67	-6.44	1
2	2	4	4	-4.70	-5.12	1.16	.35	.00A	-5.23	2
3	3	8	8	-4.67	-4.35	.93	.09	-1.24	-4.50	3
4	4	74	73	1.01	.97	.74	.75	-.55	(-3.34)	4

BSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	ZONE		50% CUM. PROBABLTY	COHERENCE M->C	C->M	ESTIM DISCR	
0	NONE		(-9.04)	-INF	-8.00		0%	0%		0
1	-7.92	.70	-6.44	-8.00	-5.66	-7.95	75%	46%	.69	1
2	-4.25	.43	-5.23	-5.66	-4.88	-5.28	6%	25%	1.32	2
3	-5.49	.43	-4.50	-4.88	-3.93	-4.96	14%	12%	1.17	3
4	-4.80	.42	(-3.34)	-3.93	+INF	-4.31	97%	93%	1.26	4

M->C = Does Measure imply Category?
C->M = Does Category imply Measure?

```
      CATEGORY PROBABILITIES: MODES - Structure measures at intersections
   -+------+------+------+------+------+------+------+------+-
1.0 +00000                                 44444444444444444444+
    |     00000                       4444                     |
 .8 +          00                   44                         +
    |            0     11111       4                           |
 .6 +             00  1     1     4                            +
 .5 +               *1       1   4                             +
 .4 +              1 00       1 4                              +
    |            11    0       *333                            |
 .2 +          11       00   ** 1  333                         +
    |     11111         22***4 22*2   3333                     |
 .0 +**********************4 00000*****************************+
   -+------+------+------+------+------+------+------+------+-
   -8     -6     -4     -2      0      2      4      6      8
      PERSON [MINUS] ITEM MEASURE
```

Item 3a: Seated Lateral Reach (Right)

ITEM DIFFICULTY MEASURE OF -2.90 ADDED TO MEASURES

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
0	0	10	10	-5.23	-5.25	1.15	1.79	NONE	(-5.73)	0
1	1	4	4	-3.78	-4.18	1.13	.60	-.83	-4.61	1
2	2	16	16	-3.73	-3.30	.92	.54	-2.26	-3.57	2
3	3	28	27	-.58	-1.13	.64	3.81	-.03	-1.33	3
4	4	44	43	2.48	2.71	1.23	1.27	3.12	(1.35)	4

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	ZONE		50% CUM. PROBABLTY	COHERENCE M->C	C->M	ESTIM DISCR	
0	NONE		(-5.73)	-INF	-5.21		100%	60%		0
1	-3.73	.49	-4.61	-5.21	-4.12	-4.75	0%	0%	.96	1
2	-5.16	.44	-3.57	-4.12	-2.73	-4.34	50%	56%	.82	2
3	-2.94	.37	-1.33	-2.73	.33	-2.85	59%	57%	1.47	3
4	.22	.35	(1.35)	.33	+INF	.26	81%	81%	.46	4

M->C = Does Measure imply Category?
C->M = Does Category imply Measure?

```
         CATEGORY PROBABILITIES: MODES - Structure measures at intersections
P       -+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-
R  1.0 +0                                                              +
O      | 000000                                                 4444444|
B   .8 +       00                                           4444       +
A      |         00                  33333333            444           |
B   .6 +           00             333         3333   44                +
I   .5 +             0   2222222333               **4                  +
L   .4 +              022      33222            44  333                +
I      |              2200   33     222      444        333            |
T   .2 +           1**1111**3          2222444             3333        +
Y      |  11111***2   333 1***11    44444422222                 3333333|
    .0 +**************44444444******************************************+
O       -+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-
F       -4    -3    -2    -1     0     1     2     3     4     5     6
         PERSON [MINUS] ITEM MEASURE
```

Item 3b: Seated Lateral Reach (Left)

ITEM DIFFICULTY MEASURE OF -2.76 ADDED TO MEASURES

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
0	0	10	10	-5.34	-5.18	.81	.80	NONE	(-5.84)	0
1	1	8	8	-3.92	-4.08	.55	.28	-1.58	-4.44	1
2	2	14	14	-3.31	-3.12	1.27	1.15	-1.44	-3.28	2
3	3	27	26	-.13	-.93	.88	4.96	-.07	-1.20	3
4	4	43	42	2.33	2.77	1.66	1.60	3.09	(1.46)	4

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	ZONE		50% CUM. PROBABLTY	COHERENCE M->C	C->M	ESTIM DISCR	
0	NONE		(-5.84)	-INF	-5.18		100%	60%		0
1	-4.34	.47	-4.44	-5.18	-3.87	-4.82	33%	50%	1.49	1
2	-4.20	.40	-3.28	-3.87	-2.48	-3.96	42%	42%	1.09	2
3	-2.82	.38	-1.20	-2.48	.44	-2.65	50%	48%	1.28	3
4	.33	.35	(1.46)	.44	+INF	.37	75%	76%	-.05	4

M->C = Does Measure imply Category?
C->M = Does Category imply Measure?

```
        CATEGORY PROBABILITIES: MODES - Structure measures at intersections
P      -+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+-
R  1.0 +                                                                                                       +
O      |0000                                                                                           44444444|
B   .8 +    000                                                                                     444        +
A      |       00                                    33333333                                  444             |
B   .6 +          00                              333           333      444                                   +
I   .5 +             00           222   33                        3**                                          +
L   .4 +                00   222     **22                        44    333                                     +
I      |                  11111***11   33        222          444            333                               |
T   .2 +      1111      22   003**1          222 444                             333                           +
Y      |11111   22222   3333000011111  44444*22222                                             33333333|
    .0 +**************4444444***************************************************************+
O      -+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+-
F      -4        -3        -2        -1         0         1         2         3         4         5         6
        PERSON [MINUS] ITEM MEASURE
```

Item 4: Pick up/touch object from the floor from a seated position

ITEM DIFFICULTY MEASURE OF -7.93 ADDED TO MEASURES

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
0	0	8	8	-6.90	-6.76	.75	.56	NONE	(-7.05)	0
1	1	4	4	-5.71	-5.72	.83	.48	2.41	-6.10	1
2	2	1	1	-2.41	-5.11	4.08	7.31	3.93	-5.50	2
3	3	25	25	-4.24*	-4.09	.83	.41	.00A	-4.49	3
4	4	64	63	1.65	1.61	.62	.89	5.09	(-1.74)	4

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	ZONE		50% CUM. PROBABLTY	COHERENCE M->C	C->M	ESTIM DISCR	
0	NONE		(-7.05)	-INF	-6.56		85%	75%		0
1	-5.52	.58	-6.10	-6.56	-5.79	-6.14	100%	25%	1.37	1
2	-4.01	.51	-5.50	-5.79	-5.15	-5.66	0%	0%	1.12	2
3	-7.93	.49	-4.49	-5.15	-2.81	-5.59	80%	68%	.95	3
4	-2.84	.38	(-1.74)	-2.81	+INF	-2.83	93%	95%	1.11	4

M->C = Does Measure imply Category?
C->M = Does Category imply Measure?

```
          CATEGORY PROBABILITIES: MODES - Structure measures at intersections
P     -+------+------+------+------+------+------+------+------+-
R  1.0 +0000000000000000000000000000                             +
O      |                           0000                     44444|
B   .8 +                               0       333        44     +
A      |                                00    3   33     4       |
B   .6 +                                 0   3      33 44        +
I   .5 +                                  0 3         *          +
L   .4 +                                   03       44 3         +
I      |                                  1*0      4    33       |
T   .2 +                               1113 *1   44       33     +
Y      |                           1111  3   0**4           33333|
    .0 +**************************************2******************+
O     -+------+------+------+------+------+------+------+------+-
F     -8     -6     -4     -2      0      2      4      6      8
          PERSON [MINUS] ITEM MEASURE
```

Item 5: Scooting forward in a chair

ITEM DIFFICULTY MEASURE OF -8.04 ADDED TO MEASURES

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
0	0	6	6	-7.13	-6.93	.88	.83	NONE	(-7.62)	0
1	1	6	6	-5.12	-5.84	1.41	2.17	1.66	-6.29	1
2	2	1	1	-4.71	-5.10	1.17	.87	4.38	-5.56	2
3	3	32	31	-3.90	-3.70	1.44	.61	.00A	-4.14	3
4	4	57	56	2.14	2.09	.69	.90	6.23	(-.71)	4

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	ZONE		50% CUM. PROBABLTY	COHERENCE M->C	C->M	ESTIM DISCR	
0	NONE		(-7.62)	-INF	-6.92		100%	66%		0
1	-6.38	.61	-6.29	-6.92	-5.91	-6.62	25%	16%	1.43	1
2	-3.66	.50	-5.56	-5.91	-5.13	-5.70	0%	0%	.73	2
3	-8.04	.48	-4.14	-5.13	-1.80	-5.62	75%	65%	.90	3
4	-1.81	.38	(-.71)	-1.80	+INF	-1.80	91%	92%	1.14	4

M->C = Does Measure imply Category?
C->M = Does Category imply Measure?

```
           CATEGORY PROBABILITIES: MODES - Structure measures at intersections
P        -+---------+---------+---------+---------+---------+---------+---------+---------+-
R     1.0 +0000000000000000000000000000000000                                             +
O         |                                 0000                    333                  4|
B      .8 +                                       00             33    33             44  +
A         |                                          00         3         33        44     |
B      .6 +                                            0       3             33   4       +
I      .5 +                                              0   3                  *         +
L      .4 +                                            11*13                 44  3        +
I         |                                           1    *11              4      33     |
T      .2 +                                       111     3  0  1      444            33  +
Y         |                                 1111         3    001*444                   3 |
       .0 +*******************************************************************************+
O        -+---------+---------+---------+---------+---------+---------+---------+---------+-
F        -8        -6        -4        -2         0         2         4         6         8
           PERSON [MINUS] ITEM MEASURE
```

Item 6: Posterior external perturbations in sitting

ITEM DIFFICULTY MEASURE OF -5.12 ADDED TO MEASURES

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
0	0	10	10	-6.66	-6.50	.98	.94	NONE	(-6.85)	0
1	1	6	6	-5.22	-5.51	1.00	.46	-.31	-5.73	1
2	2	3	3	-4.89	-5.06	.56	.13	.55	-5.17	2
3	3	7	7	-4.34	-4.43	.79	.44	-.53	-4.63	3
4	4	76	75	.81	.83	1.64	1.49	.00A	(-3.60)	4

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	ZONE		50% CUM. PROBABLTY	COHERENCE M->C	C->M	ESTIM DISCR	
0	NONE		(-6.85)	-INF	-6.26		100%	70%		0
1	-5.43	.50	-5.73	-6.26	-5.43	-5.89	50%	33%	1.53	1
2	-4.57	.46	-5.17	-5.43	-4.93	-5.27	10%	33%	.95	2
3	-5.65	.44	-4.63	-4.93	-4.14	-5.06	40%	57%	.83	3
4	-5.12	.41	(-3.60)	-4.14	+INF	-4.54	98%	92%	.90	4

M->C = Does Measure imply Category?
C->M = Does Category imply Measure?

```
        CATEGORY PROBABILITIES: MODES - Structure measures at intersections
P      -+---------+---------+---------+---------+---------+---------+---------+-
R  1.0 +0000000000000000                               4444444444444444+
O      |                  0000                   4444                  |
B   .8 +                      00               44                      +
A      |                        0             4                        |
B   .6 +                         0           4                         +
I   .5 +                          0         4                          +
L   .4 +                           0       4                           +
I      |                          111*4333                             |
T   .2 +                      111    3*1   33                          +
Y      |                  1111    2***2**2  33333                      |
    .0 +****************************4    0*****************************+
O      -+---------+---------+---------+---------+---------+---------+---------+-
F      -8        -6        -4        -2         0         2         4         6         8
        PERSON [MINUS] ITEM MEASURE
```

Item 7: Wheelchair to chair transfers

ITEM DIFFICULTY MEASURE OF -2.49 ADDED TO MEASURES

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
0	0	9	9	-4.71	-5.24	1.84	3.07	NONE	(-6.04)	0
1	1	10	10	-3.76	-4.12	2.90	5.17	-2.22	-4.48	1
2	2	10	10	-3.01	-3.15	1.40	1.59	-1.19	-3.37	2
3	3	43	42	-.71	-.44	1.45	1.31	-.98	-.74	3
4	4	30	29	3.55	3.48	.81	.71	4.39	(3.00)	4

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	---ZONE---		50% CUM. PROBABLTY	COHERENCE M->C	C->M	ESTIM DISCR	
0	NONE		(-6.04)	-INF	-5.29		33%	22%		0
1	-4.72	.48	-4.48	-5.29	-3.91	-5.00	0%	0%	-.24	1
2	-3.68	.41	-3.37	-3.91	-2.66	-3.85	15%	20%	-1.44	2
3	-3.47	.38	-.74	-2.66	1.91	-3.01	62%	65%	.27	3
4	1.90	.35	(3.00)	1.91	+INF	1.91	80%	70%	1.11	4

M->C = Does Measure imply Category?
C->M = Does Category imply Measure?

```
        CATEGORY PROBABILITIES: MODES - Structure measures at intersections
P      -+-----+------+-----+-----+-----+------+-----+-----+------+-----+-
R  1.0 +                                                                +
O      |                                33333333                        |
B   .8 +0000                        3333         33333              4444+
A      |    00                   333                  333        44     |
B   .6 +      000                 33                        333  444       +
I   .5 +         0            3                             **          +
L   .4 +        11**1111   33                            444  333       +
I      |     111    00 2***22222                      444        33     |
T   .2 + 1111      222**3  111   22222             444              3333+
Y      |1      2222 333 0000   1111   22222**44444                      |
    .0 +************4444444*********************************************+
O      -+-----+------+-----+-----+-----+------+-----+-----+------+-----+-
F      -4    -3     -2    -1     0     1      2     3     4      5     6
        PERSON [MINUS] ITEM MEASURE
```

Item 8: Support surface displacement while seated in a wheelchair

ITEM DIFFICULTY MEASURE OF -2.10 ADDED TO MEASURES

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
0	0	19	19	-4.76	-4.58	.59	.69	NONE	(-4.85)	0
1	1	11	11	-2.26	-3.41	4.17	9.90	-1.31	-3.31	1
2	2	15	15	-.64	-2.11	2.47	9.90	-.99	-1.94	2
3	3	5	5	1.48	-.54	1.04	9.90	1.85	-.85	3
4	4	52	51	1.60	2.40	6.43	6.16	.46	(.32)	4

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	---ZONE---		50% CUM. PROBABLTY	COHERENCE M->C	C->M	ESTIM DISCR	
0	NONE		(-4.85)	-INF	-4.15		92%	63%		0
1	-3.41	.38	-3.31	-4.15	-2.62	-3.82	33%	54%	1.55	1
2	-3.10	.41	-1.94	-2.62	-1.36	-2.77	25%	20%	-.50	2
3	-.26	.46	-.85	-1.36	-.22	-1.11	11%	20%	-2.85	3
4	-1.65	.43	(.32)	-.22	+INF	-.67	74%	71%	-1.41	4

M->C = Does Measure imply Category?
C->M = Does Category imply Measure?

```
       CATEGORY PROBABILITIES: MODES - Structure measures at intersections
P      -+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+-
R  1.0 +                                                                         444444444444444+
O      |000000                                                             444444               |
B   .8 +     000                                                      444                       +
A      |        00                                                 44                           |
B   .6 +          00                   22222               4                                    +
I   .5 +             00         222          22  44                                             +
L   .4 +               **1*2                 2*                                                 +
I      |          1111 2*01111             44 22                                                |
T   .2 +     11111   22    00  11114*33333**3                                                   +
Y      |111111    2222        0*******111       2***33333                                       |
    .0 +*************************44 000000*************************************************+
O      -+---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+-
F      -4        -3        -2        -1         0         1         2         3         4         5         6
          PERSON [MINUS] ITEM MEASURE
```

Item 9: Arising from a chair

ITEM DIFFICULTY MEASURE OF -.24 ADDED TO MEASURES

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
0	0	43	42	-3.66	-3.54	.21	.43	NONE	(-2.09)	0
1	1	6	6	-.95	-1.49	.67	.09	-.24	-1.03	1
2	2	4	4	-.18	-.27	.49	.18	-.19	-.38	2
3	3	13	13	.95	.87	.81	.67	-.67	.41	3
4	4	36	35	3.26	3.25	.50	.82	1.10	(2.08)	4

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	---ZONE---		50% CUM. PROBABLTY	COHERENCE M->C	C->M	ESTIM DISCR	
0	NONE		(-2.09)	-INF	-1.56		97%	95%		0
1	-.48	.44	-1.03	-1.56	-.69	-1.14	50%	50%	1.34	1
2	-.43	.51	-.38	-.69	-.05	-.63	22%	50%	1.71	2
3	-.91	.48	.41	-.05	1.25	-.26	33%	23%	1.25	3
4	.86	.36	(2.08)	1.25	+INF	1.03	86%	86%	1.05	4

M->C = Does Measure imply Category?
C->M = Does Category imply Measure?

```
       CATEGORY PROBABILITIES: MODES - Structure measures at intersections
P      -+-------+-------+-------+-------+-------+-------+-------+-------+-------+-
R  1.0 +00000                                                       444444444444+
O      |      0000000                                          4444444          |
B   .8 +             000                                   4444                 +
A      |                00                             444                      |
B   .6 +                  0                         44                          +
I   .5 +                   00         3333344                                   +
L   .4 +                     0   33   44333                                     +
I      |                   111033   44      3333                                |
T   .2 +               11111  *******            333                            +
Y      |      11111111  22***  44*****2222           33333333                   |
    .0 +*****************44444      00**************************************+
O      -+-------+-------+-------+-------+-------+-------+-------+-------+-------+-
F      -4      -3      -2      -1       0       1       2       3       4       5       6
       PERSON [MINUS] ITEM MEASURE
```

Item 10: Static standing balance

ITEM DIFFICULTY MEASURE OF -1.39 ADDED TO MEASURES

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
0	0	40	39	-4.94	-4.90	.21	.49	NONE	(-3.19)	0
1	1	5	5	-2.21	-2.87	2.58	1.82	-.37	-2.18	1
2	2	1	1	-.56	-1.61	1.50	.79	.80	-1.66	2
3	3	6	6	-.86*	-.45	1.34	.35	-1.46	-1.13	3
4	4	50	49	2.67	2.67	1.66	1.05	.00A	(-.03)	4

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	---ZONE---		50% CUM. PROBABLTY	COHERENCE M->C	C->M	ESTIM DISCR	
0	NONE		(-3.19)	-INF	-2.65		97%	97%		0
1	-1.76	.49	-2.18	-2.65	-1.90	-2.26	50%	40%	1.13	1
2	-.59	.68	-1.66	-1.90	-1.42	-1.73	0%	0%	.67	2
3	-2.85	.67	-1.13	-1.42	-.63	-1.61	33%	33%	.53	3
4	-1.39	.47	(-.03)	-.63	+INF	-.99	91%	90%	.99	4

M->C = Does Measure imply Category?
C->M = Does Category imply Measure?

```
          CATEGORY PROBABILITIES: MODES - Structure measures at intersections
P      -+-------+-------+-------+-------+-------+-------+-------+-------+-
R  1.0 +0000000000000000                                 4444444444444444+
O      |                 0000                       4444                 |
B   .8 +                     00                 44                       +
A      |                       0            44                           |
B   .6 +                        0         4                              +
I   .5 +                         0      4                                +
L   .4 +                          0 4                                    +
I      |                        11*3*333                                 |
T   .2 +                     111  3*1    333                             +
Y      |                 1111    3**2**1     3333                        |
    .0 +**************************4     *********************************+
O      -+-------+-------+-------+-------+-------+-------+-------+-------+-
F      -8      -6      -4      -2       0       2       4       6       8
          PERSON [MINUS] ITEM MEASURE
```

Item 11: Stand to sit

ITEM DIFFICULTY MEASURE OF -.57 ADDED TO MEASURES

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
0	0	43	42	-3.65	-3.56	.23	.40	NONE	(-2.01)	0
1	1	5	5	-.76	-1.66	1.69	.19	.18	-1.14	1
2	2	1	1	-2.08*	-.57	2.35	2.26	1.11	-.65	2
3	3	10	10	.08	.44	.70	.23	-1.83	-.10	3
4	4	43	42	2.97	2.87	.31	.61	.55	(1.20)	4

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	ZONE		50% CUM. PROBABLTY	COHERENCE M->C	C->M	ESTIM DISCR	
0	NONE		(-2.01)	-INF	-1.56		95%	93%		0
1	-.39	.46	-1.14	-1.56	-.88	-1.12	25%	20%	1.24	1
2	.54	.57	-.65	-.88	-.41	-.74	0%	0%	.87	2
3	-2.40	.56	-.10	-.41	.48	-.66	42%	30%	1.05	3
4	-.02	.40	(1.20)	.48	+INF	.19	93%	95%	1.35	4

M->C = Does Measure imply Category?
C->M = Does Category imply Measure?

```
        CATEGORY PROBABILITIES: MODES - Structure measures at intersections
P      -+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-
R  1.0 +00000000                                      444444444444444+
O      |        000000                          4444444              |
B   .8 +              000                   444                      +
A      |                 00              444                         |
B   .6 +                   00          44                            +
I   .5 +                     0       44                              +
L   .4 +                      0 333*333                              +
I      |                       * 44    333                           |
T   .2 +              111111**1*          3333                       +
Y      |       1111111   3334**2***1            3333333              |
    .0 +**********************2    22************************************+
O      -+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-
F      -4    -3    -2    -1     0     1     2     3     4     5     6
        PERSON [MINUS] ITEM MEASURE
```

Item 12: Static standing balance with eyes closed

ITEM DIFFICULTY MEASURE OF .37 ADDED TO MEASURES

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
0	0	48	47	-4.37	-4.41	2.42	9.90	NONE	(-2.10)	0
1	1	4	4	-1.42	-1.36	.84	.17	-.63	-1.00	1
2	2	7	7	.08	-.11	1.05	.58	-1.59	-.24	2
3	3	7	7	.40	.96	1.39	.25	.02	.60	3
4	4	36	35	3.63	3.61	.68	.81	.00A	(1.91)	4

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	ZONE		50% CUM. PROBABLTY	COHERENCE M->C	C->M	ESTIM DISCR	
0	NONE		(-2.10)	-INF	-1.57		95%	91%		0
1	-.26	.50	-1.00	-1.57	-.60	-1.13	14%	25%	.85	1
2	-1.22	.53	-.24	-.60	.14	-.67	40%	28%	.52	2
3	.40	.49	.60	.14	1.28	.16	30%	42%	.04	3
4	.37	.43	(1.91)	1.28	+INF	.90	91%	86%	1.09	4

M->C = Does Measure imply Category?
C->M = Does Category imply Measure?

```
          CATEGORY PROBABILITIES: MODES - Structure measures at intersections
P      -+---------+---------+---------+---------+---------+---------+---------+---------+-
R  1.0 +0000000000000000                                    4444444444444444+
O      |                 000                        4444                    |
B   .8 +                    00                    44                        +
A      |                      0                  4                          |
B   .6 +                       0               4                            +
I   .5 +                        0             4                             +
L   .4 +                        0 222 4                                     +
I      |                          *  3*333                                  |
T   .2 +                     1**1*34 22 33                                  +
Y      |                 1111*2 334**1  22 33333                            |
    .0 +*************************44     0*******************************+
O      -+---------+---------+---------+---------+---------+---------+---------+---------+-
F      -8        -6        -4        -2         0         2         4         6         8
          PERSON [MINUS] ITEM MEASURE
```

Item 13: Static standing balance with feet together and eyes open

ITEM DIFFICULTY MEASURE OF -.18 ADDED TO MEASURES

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
0	0	48	47	-3.40	-3.33	.34	.40	NONE	(-1.53)	0
1	1	2	2	-.19	-1.19	1.17	.04	1.20	-.73	1
2	2	5	5	.21	-.17	.66	1.01	-1.37	-.23	2
3	3	5	5	-.08*	.77	1.66	.16	.44	.31	3
4	4	42	41	3.06	2.98	.19	.45	-.28	(1.32)	4

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	ZONE		50% CUM. PROBABLTY	COHERENCE M->C	C->M	ESTIM DISCR	
0	NONE		(-1.53)	-INF	-1.14		100%	95%		0
1	1.02	.49	-.73	-1.14	-.46	-.64	0%	0%	1.16	1
2	-1.55	.53	-.23	-.46	.01	-.49	50%	80%	1.22	2
3	.26	.51	.31	.01	.80	-.04	50%	40%	.69	3
4	-.46	.44	(1.32)	.80	+INF	.37	97%	95%	1.24	4

M->C = Does Measure imply Category?
C->M = Does Category imply Measure?

```
        CATEGORY PROBABILITIES: MODES - Structure measures at intersections
P      -+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-
R  1.0 +000000000000                             44444444444444444444+
O      |            0000                   444444                    |
B   .8 +                000             444                          +
A      |                   0          44                             |
B   .6 +                    0       44                               +
I   .5 +                     0     4                                 +
L   .4 +                      0   4                                  +
I      |                      2*2*                                   |
T   .2 +                   222 4*****3333                            +
Y      |              1****1****1100 22223333333                     |
    .0 +********************4     11*********************************+
O      -+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-
F     -4    -3    -2    -1     0     1     2     3     4     5     6
        PERSON [MINUS] ITEM MEASURE
```

Item 14: External perturbations in standing

ITEM DIFFICULTY MEASURE OF -1.78 ADDED TO MEASURES

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
0	0	48	47	-4.48	-4.43	.52	.49	NONE	(-1.99)	0
1	1	3	3	-.15	-1.55	2.78	.97	1.70	-1.15	1
2	2	2	2	.82	-.40	1.94	1.33	1.28	-.60	2
3	3	13	13	.20*	.68	1.18	.30	.00A	.13	3
4	4	36	35	3.62	3.56	.61	.84	2.59	(1.97)	4

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	ZONE		50% CUM. PROBABLTY	COHERENCE M->C	C->M	ESTIM DISCR	
0	NONE		(-1.99)	-INF	-1.57		97%	93%		0
1	-.08	.53	-1.15	-1.57	-.86	-1.10	0%	0%	1.20	1
2	-.50	.59	-.60	-.86	-.30	-.80	0%	0%	-.05	2
3	-1.78	.55	.13	-.30	1.05	-.61	45%	38%	.51	3
4	.81	.40	(1.97)	1.05	+INF	.90	91%	88%	1.21	4

M->C = Does Measure imply Category?
C->M = Does Category imply Measure?

```
        CATEGORY PROBABILITIES: MODES - Structure measures at intersections
P      -+-------+-------+-------+-------+-------+-------+-------+-------+-
R  1.0 +0000000000000000000000000000                             444444444+
O      |                            000                    4444           |
B   .8 +                               00              444                +
A      |                                  0           4                   |
B   .6 +                                   0     3    4                   +
I   .5 +                                   0  33 3*4                      +
L   .4 +                                    03    4 33                    +
I      |                                    *    4    3                   |
T   .2 +                                 1*1044        333                +
Y      |                            11111*22***22          3333           |
    .0 +*********************************444    **************************+
O      -+-------+-------+-------+-------+-------+-------+-------+-------+-
F      -8      -6      -4      -2       0       2       4       6       8
        PERSON [MINUS] ITEM MEASURE
```

Item 15: Standing forward reach

ITEM DIFFICULTY MEASURE OF .29 ADDED TO MEASURES

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
0	0	45	44	-3.54	-3.48	.38	.39	NONE	(-1.96)	0
1	1	2	2	-1.25	-1.39	.81	.11	.45	-1.06	1
2	2	10	10	-.08	.01	.41	.18	-2.56	-.29	2
3	3	23	23	1.72	1.54	.75	.94	-.40	1.39	3
4	4	22	22	4.05	4.08	.89	.98	2.60	(3.93)	4

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	---ZONE---		50% CUM. PROBABLTY	COHERENCE M->C	C->M	ESTIM DISCR	
0	NONE		(-1.96)	-INF	-1.54		97%	91%		0
1	.74	.48	-1.06	-1.54	-.69	-1.03	0%	0%	1.12	1
2	-2.27	.50	-.29	-.69	.29	-.87	70%	70%	1.23	2
3	-.11	.41	1.39	.29	2.93	.08	66%	69%	1.41	3
4	2.79	.39	(3.93)	2.93	+INF	2.84	80%	72%	.95	4

M->C = Does Measure imply Category?
C->M = Does Category imply Measure?

```
        CATEGORY PROBABILITIES: MODES - Structure measures at intersections
P      -+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-
R  1.0 +000000                                                                    4444+
O      |        00000                                                      4444444    |
B   .8 +             00                                                4444           +
A      |               00                333333                     444               |
B   .6 +                 0           333         3333   44                            +
I   .5 +                  0        33               3*4                               +
L   .4 +                   0222**22              444 333                              +
I      |                  220*3     222        44        333                          |
T   .2 +               222 33 0       22*444                  3333                    +
Y      |           ***1***1111*00 44444 222222                     3333333            |
    .0 +**************44444444*******************************************************+
O      -+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-
F      -4      -3      -2      -1       0       1       2       3       4       5       6
        PERSON [MINUS] ITEM MEASURE
```

Item 16: Pick up/touch object from the floor from a standing position

ITEM DIFFICULTY MEASURE OF .98 ADDED TO MEASURES

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
0	0	51	50	-4.18	-4.18	2.06	7.46	NONE	(-1.62)	0
1	1	7	7	-.14	-.82	1.85	2.49	-1.27	-.40	1
2	2	2	2	-.05	.31	.33	.10	.09	.22	2
3	3	9	9	.79	1.26	1.35	.36	-1.73	.88	3
4	4	33	32	3.82	3.82	.53	.74	.00A	(2.25)	4

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	----ZONE----		50% CUM. PROBABLTY	COHERENCE M->C	C->M	ESTIM DISCR	
0	NONE		(-1.62)	-INF	-.98		95%	90%		0
1	-.28	.46	-.40	-.98	-.06	-.64	33%	42%	.99	1
2	1.07	.55	.22	-.06	.51	.11	33%	50%	.04	2
3	-.74	.54	.88	.51	1.52	.31	50%	55%	.40	3
4	.98	.43	(2.25)	1.52	+INF	1.23	87%	84%	1.02	4

M->C = Does Measure imply Category?
C->M = Does Category imply Measure?

```
          CATEGORY PROBABILITIES: MODES - Structure measures at intersections
P      -+-------+-------+-------+-------+-------+-------+-------+-------+-
R  1.0 +00000000000000                             444444444444444444+
O      |              0000                   4444                       |
B   .8 +                  00               44                           +
A      |                    0            44                             |
B   .6 +                     0          4                               +
I   .5 +                      0        4                                +
L   .4 +                       *   33*3                                 +
I      |                    111 ** 4  33                                |
T   .2 +                  11    30*     333                             +
Y      |              1111    *******2      3333                        |
    .0 +*************************44    0***********************************+
O      -+-------+-------+-------+-------+-------+-------+-------+-------+-
F      -8      -6      -4      -2       0       2       4       6       8
          PERSON [MINUS] ITEM MEASURE
```

Item 17: Standing trunk rotation

ITEM DIFFICULTY MEASURE OF -.99 ADDED TO MEASURES

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
0	0	44	43	-4.76	-4.62	.22	.45	NONE	(-2.77)	0
1	1	8	8	-1.10	-1.56	.40	.08	-.34	-1.25	1
2	2	11	11	.64	.03	1.57	5.10	.00A	-.03	2
3	3	5	5	.99	1.19	.40	.11	2.37	.99	3
4	4	34	33	3.68	3.78	1.20	1.18	1.31	(2.17)	4

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	ZONE		50% CUM. PROBABLTY	COHERENCE M->C	C->M	ESTIM DISCR	
0	NONE		(-2.77)	-INF	-2.06		97%	97%		0
1	-1.33	.45	-1.25	-2.06	-.61	-1.73	55%	62%	1.43	1
2	-.99	.48	-.03	-.61	.49	-.71	42%	27%	1.51	2
3	1.38	.47	.99	.49	1.62	.68	23%	60%	.19	3
4	.32	.45	(2.17)	1.62	+INF	1.18	86%	73%	.71	4

M->C = Does Measure imply Category?
C->M = Does Category imply Measure?

```
        CATEGORY PROBABILITIES: MODES - Structure measures at intersections
P      -+--------+--------+--------+--------+--------+--------+--------+--------+-
R  1.0 +0000000000000000                                     4444444444444+
O      |                0000                         4444                  |
B   .8 +                    00                           4                     +
A      |                      0                       44                   |
B   .6 +                       0                     4                     +
I   .5 +                        0      2222  4                     +
L   .4 +                          01*2      24                     +
I      |                        111*011     42                     |
T   .2 +                    111   2   0 1**33**3                   +
Y      |               1111  222     3**411     2*333              |
    .0 +*********************************44 000***********************+
O      -+--------+--------+--------+--------+--------+--------+--------+--------+-
F      -8       -6       -4       -2        0        2        4        6        8
        PERSON [MINUS] ITEM MEASURE
```

Item 18: Turn 180 degrees

ITEM DIFFICULTY MEASURE OF 1.52 ADDED TO MEASURES

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
0	0	59	58	-2.77	-2.70	.44	.56	NONE	(-.39)	0
1	1	5	5	1.09	.38	1.17	.09	.02	.72	1
2	2	10	10	1.30	1.45	.60	.30	-1.23	1.50	2
3	3	5	5	2.20	2.36	.38	.10	1.03	2.31	3
4	4	23	23	4.34	4.21	.60	.66	.17	(3.48)	4

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	---ZONE---		50% CUM. PROBABLTY	COHERENCE M->C	C->M	ESTIM DISCR	
0	NONE		(-.39)	-INF	.14		98%	94%		0
1	1.54	.43	.72	.14	1.13	.59	33%	60%	1.25	1
2	.29	.46	1.50	1.13	1.88	1.01	50%	50%	1.12	2
3	2.56	.46	2.31	1.88	2.92	1.98	16%	20%	1.47	3
4	1.69	.47	(3.48)	2.92	+INF	2.49	95%	82%	1.33	4

M->C = Does Measure imply Category?
C->M = Does Category imply Measure?

```
        CATEGORY PROBABILITIES: MODES - Structure measures at intersections
P      -+---------+---------+---------+---------+---------+---------+---------+---------+---------+-
R  1.0 +000000                                                        4444444444444444+
O      |         000000                                          444444               |
B   .8 +               000                                    444                     +
A      |                  00                                44                        |
B   .6 +                    0                             44                          +
I   .5 +                     00                         4                             +
L   .4 +                        0222222244                                            +
I      |                       220       422                                          |
T   .2 +                  1***111**3**333**333                                        +
Y      |          11111***2      33*****1      222**333333                            |
    .0 +*********************44      0*****************************************+
O      -+---------+---------+---------+---------+---------+---------+---------+---------+---------+-
F      -4        -3        -2        -1         0         1         2         3         4         5         6
        PERSON [MINUS] ITEM MEASURE
```

Item 19: Alternating step test

ITEM DIFFICULTY MEASURE OF 3.04 ADDED TO MEASURES

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
0	0	64	63	-2.46	-2.38	.75	.83	NONE	(.08)	0
1	1	11	11	1.60	1.17	.87	.19	-1.57	1.59	1
2	2	9	9	2.23	2.56	1.00	.67	-.97	2.75	2
3	3	12	12	4.36	4.22	.39	.42	.06	4.40	3
4	4	6	6	5.51	5.32	.73	.83	2.47	(6.68)	4

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	----ZONE----		50% CUM. PROBABLTY	COHERENCE M->C	C->M	ESTIM DISCR	
0	NONE		(.08)	-INF	.80		96%	92%		0
1	1.48	.38	1.59	.80	2.17	1.11	41%	63%	1.28	1
2	2.07	.44	2.75	2.17	3.45	2.15	28%	22%	.93	2
3	3.11	.49	4.40	3.45	5.72	3.30	69%	75%	1.48	3
4	5.51	.54	(6.68)	5.72	+INF	5.60	75%	50%	1.34	4

M->C = Does Measure imply Category?
C->M = Does Category imply Measure?

```
         CATEGORY PROBABILITIES: MODES - Structure measures at intersections
P      -+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-
R  1.0 +                                                                           4444+
O      |0000                                                                 4444444   |
B   .8 +    0000                                                        4444           +
A      |        00                                                  444                |
B   .6 +          00                               33333333    44                      +
I   .5 +            00                         333        33*4                         +
L   .4 +              1**11 22222**2                     444 333                       +
I      |          1111   *2*11133   222     44          333                            |
T   .2 +      1111    222 00*33111       2**4                 3333                     +
Y      |11111     2222    333 000  1***44  222222                     3333333          |
    .0 +******************4444444***000********************************************+
O      -+-------+-------+-------+-------+-------+-------+-------+-------+-------+-------+-
F      -4      -3      -2      -1       0       1       2       3       4       5       6
        PERSON [MINUS] ITEM MEASURE
```

Item 20: Balance in tandem/stride stance

ITEM DIFFICULTY MEASURE OF 2.48 ADDED TO MEASURES

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
0	0	52	51	-4.18	-4.09	.80	.94	NONE	(-1.66)	0
1	1	10	10	.28	-.15	.78	.13	-2.72	-.06	1
2	2	17	17	1.24	1.43	.52	.33	-2.30	1.69	2
3	3	4	4	4.87	2.96	1.00	3.46	1.11	3.09	3
4	4	19	19	4.89	5.10	1.03	.87	.00A	(4.34)	4

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	ZONE		50% CUM. PROBABLTY	COHERENCE M->C	C->M	ESTIM DISCR	
0	NONE		(-1.66)	-INF	-.94		100%	96%		0
1	-.24	.43	-.06	-.94	.73	-.62	53%	70%	1.33	1
2	.18	.43	1.69	.73	2.48	.51	73%	64%	1.03	2
3	3.59	.51	3.09	2.48	3.77	2.80	14%	25%	1.41	3
4	2.48	.54	(4.34)	3.77	+INF	3.33	82%	73%	.50	4

M->C = Does Measure imply Category?
C->M = Does Category imply Measure?

```
        CATEGORY PROBABILITIES: MODES - Structure measures at intersections
P      -+---------+---------+---------+---------+---------+---------+---------+---------+-
R  1.0 +000000000                                               444444444444444444+
O      |          0000                                      444                  |
B   .8 +              00                                 44                      +
A      |                0         222                 4                          |
B   .6 +                 0       2    22           4                             +
I   .5 +                  0    22       2 4                                      +
L   .4 +                    *1*            *                                     +
I      |                111 * 111       4 2                                      |
T   .2 +             11    2 00   11  3*333*3                                    +
Y      |          1111  222    0033**1     2*3333                                |
    .0 +**************************000*************************************+
O      -+---------+---------+---------+---------+---------+---------+---------+---------+-
F      -8        -6        -4        -2         0         2         4         6         8
        PERSON [MINUS] ITEM MEASURE
```

Item 21a: Single leg stance on right lower extremity

ITEM DIFFICULTY MEASURE OF 6.57 ADDED TO MEASURES

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
0	0	62	61	-3.43	-3.37	.82	.92	NONE	(-.53)	0
1	1	27	26	2.17	1.97	1.38	.37	-6.01	3.15	1
2	2	1	1	4.32	4.35	.20	.06	.00A	4.66	2
3	3	3	3	5.25	5.39	2.01	1.51	-2.73	5.29	3
4	4	9	9	5.84	6.00	2.11	5.81	-1.96	(6.23)	4

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	---ZONE---		50% CUM. PROBABLTY	COHERENCE M->C	C->M	ESTIM DISCR	
0	NONE		(-.53)	-INF	.56		96%	93%		0
1	.57	.36	3.15	.56	4.26	.57	76%	74%	1.19	1
2	6.57	.66	4.66	4.26	4.97	4.75	25%	100%	.84	2
3	3.84	.68	5.29	4.97	5.75	4.88	0%	0%	.48	3
4	4.61	.60	(6.23)	5.75	+INF	5.32	75%	66%	.60	4

M->C = Does Measure imply Category?
C->M = Does Category imply Measure?

```
        CATEGORY PROBABILITIES: MODES - Structure measures at intersections
P     -+------+------+------+------+------+------+------+------+-
R  1.0 +                                444444444444444444444444+
O      |0             11111         4444                        |
B   .8 + 000       111     1      44                            +
A      |    0     1         1    4                              |
B   .6 +     00 11           1  4                               +
I   .5 +       *             1 4                                +
L   .4 +     11 00            14                                +
I      |    1     0           41                                |
T   .2 + 111       00       *33*333                             +
Y      |1            00000 **222 1  3333                        |
    .0 ********************0000********************************+
O     -+------+------+------+------+------+------+------+------+-
F     -8     -6     -4     -2      0      2      4      6      8
        PERSON [MINUS] ITEM MEASURE
```

Item 21b: Single leg stance on left lower extremity

ITEM DIFFICULTY MEASURE OF 5.42 ADDED TO MEASURES

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
0	0	61	60	-3.56	-3.43	.52	.51	NONE	(-.66)	0
1	1	27	26	2.22	1.87	.62	.25	-4.97	2.81	1
2	2	3	3	3.57	4.35	2.03	2.38	.00A	4.70	2
3	3	3	3	5.33	5.46	.28	.09	-.43	5.44	3
4	4	8	8	6.18	6.06	.56	.54	-.62	(6.49)	4

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	ZONE		50% CUM. PROBABLTY	COHERENCE M->C	C->M	ESTIM DISCR	
0	NONE		(-.66)	-INF	.43		96%	93%		0
1	.45	.37	2.81	.43	4.20	.44	77%	77%	1.42	1
2	5.42	.59	4.70	4.20	5.07	4.61	0%	0%	.80	2
3	4.99	.64	5.44	5.07	5.97	5.05	50%	100%	1.38	3
4	4.80	.60	(6.49)	5.97	+INF	5.54	83%	62%	1.28	4

M->C = Does Measure imply Category?
C->M = Does Category imply Measure?

```
          CATEGORY PROBABILITIES: MODES - Structure measures at intersections
P      -+---------+---------+---------+---------+---------+---------+---------+---------+-
R  1.0 +0                                                  4444444444444444444+
O      | 0000                                          4444                   |
B   .8 +     00            11111111             44                            +
A      |       00     11           1          4                               |
B   .6 +         0  1                1      4                                 +
I   .5 +          **                  1   4                                   +
L   .4 +         1  0                 1 4                                     +
I      |       11     00                *                                     |
T   .2 +     11          00         222***333                                 +
Y      | 1111              00**223**    *2 33333                              |
    .0 +*************************000000***********************************+
       -+---------+---------+---------+---------+---------+---------+---------+---------+-
O
F      -8        -6        -4        -2         0         2         4         6         8
          PERSON [MINUS] ITEM MEASURE
```

Item 22: Walking over level surface

ITEM DIFFICULTY MEASURE OF 1.00 ADDED TO MEASURES

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
0	0	46	45	-3.51	-3.38	.27	.37	NONE	(-1.88)	0
1	1	6	6	.04	-.86	.87	.17	-1.02	-.65	1
2	2	19	19	.55	.81	.61	.33	-2.13	.65	2
3	3	14	14	2.31	2.40	.21	.14	.86	2.61	3
4	4	17	17	4.96	4.54	.19	.35	2.29	(4.55)	4

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	---ZONE---		50% CUM. PROBABLTY	COHERENCE M->C	C->M	ESTIM DISCR	
0	NONE		(-1.88)	-INF	-1.32		100%	95%		0
1	-.02	.42	-.65	-1.32	-.07	-.88	25%	50%	1.36	1
2	-1.13	.44	.65	-.07	1.64	-.36	76%	52%	1.08	2
3	1.87	.39	2.61	1.64	3.70	1.74	87%	100%	1.79	3
4	3.30	.44	(4.55)	3.70	+INF	3.48	100%	100%	1.79	4

M->C = Does Measure imply Category?
C->M = Does Category imply Measure?

```
        CATEGORY PROBABILITIES: MODES - Structure measures at intersections
P      -+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-
R  1.0 +                                                         44444+
O      |000000                                           4444444      |
B   .8 +      000                                    4444             +
A      |         00                                44                 |
B   .6 +           0      222222222              444                  +
I   .5 +            00  22          22 3333333344                     +
L   .4 +              02            33*22    443333                   +
I      |             2200         333     2244       333              |
T   .2 +         11**1111**1 333         444222        3333           +
Y      | 111111**22     333****111 4444      2222           3333333   |
    .0 +****************444444****************************************+
O      -+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+-
F      -4    -3    -2    -1     0     1     2     3     4     5     6
        PERSON [MINUS] ITEM MEASURE
```

Item 23: Walking with horizontal head turns

ITEM DIFFICULTY MEASURE OF 4.01 ADDED TO MEASURES

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
0	0	71	70	-2.95	-2.89	.33	.48	NONE	(.76)	0
1	1	4	4	1.91	1.43	.08	.00	-1.23	1.89	1
2	2	9	9	2.79	2.68	.45	.21	-2.75	2.89	2
3	3	5	5	4.88	4.23	.45	1.59	.00A	4.09	3
4	4	13	13	5.48	5.62	2.03	1.56	.02	(5.54)	4

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	ZONE		50% CUM. PROBABLTY	COHERENCE M->C	C->M	ESTIM DISCR	
0	NONE		(.76)	-INF	1.28		100%	95%		0
1	2.78	.45	1.89	1.28	2.38	1.74	40%	100%	1.27	1
2	1.26	.50	2.89	2.38	3.47	2.15	85%	66%	1.50	2
3	4.01	.58	4.09	3.47	4.87	3.62	50%	40%	1.33	3
4	4.03	.55	(5.54)	4.87	+INF	4.52	84%	84%	.69	4

M->C = Does Measure imply Category?
C->M = Does Category imply Measure?

```
       CATEGORY PROBABILITIES: MODES - Structure measures at intersections
P      -+-------+-------+-------+-------+-------+-------+-------+-------+-
R  1.0 +000000000000000                             44444444444444444+
O      |               0000                      4444                |
B   .8 +                   0                   44                    +
A      |                    0                 4                      |
B   .6 +                     0    2          4                       +
I   .5 +                      0 22 22       4                        +
L   .4 +                       *      2 4                            +
I      |                      2 0    33*333                          |
T   .2 +                     *11103344 22 333                        +
Y      |               111**  33***1     22  3333                    |
    .0 +***********************444   0**********************************+
O      -+-------+-------+-------+-------+-------+-------+-------+-------+-
F      -8      -6      -4      -2       0       2       4       6       8
       PERSON [MINUS] ITEM MEASURE
```

Item 24: Walking with change in direction

ITEM DIFFICULTY MEASURE OF 3.41 ADDED TO MEASURES

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
0	0	72	71	-2.89	-2.86	.42	.44	NONE	(.82)	0
1	1	0	0			.00	.00	NULL	1.69	1
2	2	10	10	2.32	2.44	.40	.06	-3.12	2.52	2
3	3	8	8	4.28	4.07	.34	.45	.00A	3.98	3
4	4	12	12	5.78	5.63	.48	.67	1.11	(5.81)	4

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.
Unobserved category. Consider: STKEEP=NO

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	ZONE		50% CUM. PROBABLTY	COHERENCE M->C	C->M	ESTIM DISCR	
0	NONE		(.82)	-INF	1.21		100%	94%		0
1	NULL		1.69	1.21	2.08	1.76	0%	0%	1.00	1
2	.29	.52	2.52	2.08	3.15	1.76	87%	70%	1.27	2
3	3.41	.53	3.98	3.15	5.00	3.23	71%	62%	1.75	3
4	4.52	.52	(5.81)	5.00	+INF	4.75	83%	83%	1.36	4

M->C = Does Measure imply Category?
C->M = Does Category imply Measure?

```
          CATEGORY PROBABILITIES: MODES - Structure measures at intersections
P         -+------+------+------+------+------+------+------+------+-
R   1.0   +00000000000000000                        44444444444444+
O         |                 00                  4444              |
B    .8   +                   0               44                  +
A         |                    0            44                    |
B    .6   +                    0  222       4                     +
I    .5   +                     02   2 333 4                      +
L    .4   +                     20  3*  4*3                       +
I         |                    2 0  3  2*  33                     |
T    .2   +                   2   *3   4 2   33                   +
Y         |                 22  33 0*44  222   33333              |
     .0   +*************************]*****************************+
O         -+------+------+------+------+------+------+------+------+-
F         -8     -6     -4     -2      0      2      4      6      8
          PERSON [MINUS] ITEM MEASURE
```

Item 25: Stepping over object while walking

ITEM DIFFICULTY MEASURE OF 2.14 ADDED TO MEASURES

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
0	0	73	72	-2.82	-2.80	.58	.52	NONE	(.99)	0
1	1	0	0			.00	.00	NULL	1.72	1
2	2	5	5	2.80	2.23	.67	.18	.07	2.31	2
3	3	10	10	3.29	3.60	1.04	.52	.00A	3.32	3
4	4	14	14	5.56	5.43	1.31	.84	2.08	(5.41)	4

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.
Unobserved category. Consider: STKEEP=NO

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	ZONE		50% CUM. PROBABLTY	COHERENCE M->C	C->M	ESTIM DISCR	
0	NONE		(.99)	-INF	1.33		100%	93%		0
1	NULL		1.72	1.33	2.02	1.87	0%	0%	1.00	1
2	2.21	.55	2.31	2.02	2.68	1.87	50%	40%	1.59	2
3	2.14	.54	3.32	2.68	4.49	2.45	77%	70%	.58	3
4	4.23	.50	(5.41)	4.49	+INF	4.33	92%	92%	1.27	4

M->C = Does Measure imply Category?
C->M = Does Category imply Measure?

```
        CATEGORY PROBABILITIES: MODES - Structure measures at intersections
P      -+-------+-------+-------+-------+-------+-------+-------+-------+-
R  1.0 +000000000000000000000000                         44444444444+
O      |                        00                   4444           |
B   .8 +                          0                44               +
A      |                           0             44                 |
B   .6 +                           0     333   4                    +
I   .5 +                            0  3   334                      +
L   .4 +                            0 3    4433                     +
I      |                            2*22  4    33                   |
T   .2 +                           23 0 **       33                 +
Y      |                         22*3  4*  222     3333             |
    .0 +********************************11*****************************+
O      -+-------+-------+-------+-------+-------+-------+-------+-------+-
F      -8      -6      -4      -2       0       2       4       6       8
        PERSON [MINUS] ITEM MEASURE
```

Item 26: Walking while carrying an object with 2 hands

ITEM DIFFICULTY MEASURE OF 3.44 ADDED TO MEASURES

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
0	0	72	71	-2.89	-2.83	.29	.47	NONE	(.91)	0
1	1	4	4	2.18	1.50	.69	.03	-.57	2.00	1
2	2	7	7	2.66	2.71	.44	.16	-1.89	2.88	2
3	3	7	7	4.52	4.25	.96	1.34	.00A	4.06	3
4	4	12	12	5.65	5.65	.55	.78	1.03	(5.78)	4

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	ZONE		50% CUM. PROBABLTY	COHERENCE M->C	C->M	ESTIM DISCR	
0	NONE		(.91)	-INF	1.42		100%	94%		0
1	2.87	.46	2.00	1.42	2.44	1.88	20%	50%	1.27	1
2	1.56	.52	2.88	2.44	3.40	2.29	60%	42%	1.41	2
3	3.44	.57	4.06	3.40	4.99	3.38	57%	57%	1.52	3
4	4.47	.53	(5.78)	4.99	+INF	4.73	83%	83%	1.00	4

M->C = Does Measure imply Category?
C->M = Does Category imply Measure?

```
         CATEGORY PROBABILITIES: MODES - Structure measures at intersections
P      -+---------+---------+---------+---------+---------+---------+---------+---------+-
R   1.0 +0000000000000000000000                                   4444444444444444444444+
O       |                      000                           4444                       |
B    .8 +                          00                     44                            +
A       |                             0                  4                              |
B    .6 +                              0                4                               +
I    .5 +                              0     2      44                                  +
L    .4 +                               022 2*33*33                                     +
I       |                               20 33 24       33                               |
T    .2 +                             1*11*     422       33                            +
Y       |                      111**2 33 ***        222     3333                        |
     .0 +*********************************444  0****************************************+
O      -+---------+---------+---------+---------+---------+---------+---------+---------+-
F      -8        -6        -4        -2         0         2         4         6         8
         PERSON [MINUS] ITEM MEASURE
```

Item 27: Walking up/down stairs

ITEM DIFFICULTY MEASURE OF 2.43 ADDED TO MEASURES

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
0	0	53	52	-4.11	-4.04	.61	.60	NONE	(-1.49)	0
1	1	8	8	.01	-.17	.55	.08	-2.39	-.06	1
2	2	16	16	1.89	1.47	1.13	1.23	-2.44	1.41	2
3	3	18	18	3.58	3.88	1.16	1.07	.00A	4.21	3
4	4	7	7	5.98	5.84	.78	.82	3.54	(7.09)	4

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	ZONE		50% CUM. PROBABLTY	COHERENCE M->C	C->M	ESTIM DISCR	
0	NONE		(-1.49)	-INF	-.84		96%	90%		0
1	.04	.44	-.06	-.84	.60	-.47	40%	50%	1.28	1
2	-.01	.44	1.41	.60	2.56	.40	50%	56%	1.32	2
3	2.43	.43	4.21	2.56	6.05	2.49	61%	61%	.50	3
4	5.97	.52	(7.09)	6.05	+INF	6.00	66%	57%	1.20	4

M->C = Does Measure imply Category?
C->M = Does Category imply Measure?

```
        CATEGORY PROBABILITIES: MODES - Structure measures at intersections
P      -+------+------+------+------+------+------+------+------+-
R  1.0 +0000000000                                         444444+
O      |          0000                                 4444      |
B   .8 +              0                              44          +
A      |               00              3333333     44            |
B   .6 +                 0     2222   3       33  4              +
I   .5 +                  0  22    2*3          **               +
L   .4 +                   02      3 2         4  3              +
I      |                 11**11  33   22     44    33            |
T   .2 +              111 2  0 1*       22 44        33          +
Y      |          1111 222   3** 111    44*222         3333      |
    .0 +*********************444*********************************+
O      -+------+------+------+------+------+------+------+------+-
F      -8     -6     -4     -2      0      2      4      6      8
        PERSON [MINUS] ITEM MEASURE
```

Item 28: Walking up/down an incline

ITEM DIFFICULTY MEASURE OF 4.15 ADDED TO MEASURES

CATEGORY LABEL	SCORE	OBSERVED COUNT	%	OBSVD AVRGE	SAMPLE EXPECT	INFIT MNSQ	OUTFIT MNSQ	STRUCTURE CALIBRATN	CATEGORY MEASURE	
0	0	78	77	-2.51	-2.44	.23	.58	NONE	(1.83)	0
1	1	7	7	3.22	2.48	.13	.01	-.99	3.35	1
2	2	3	3	3.83	4.06	.95	.87	.00A	4.19	2
3	3	3	3	5.46	5.20	.33	.28	.52	4.91	3
4	4	10	10	5.97	5.93	.93	.83	.23	(6.04)	4
MISSING		1	1	4.56						

OBSERVED AVERAGE is mean of measures in category. It is not a parameter estimate.

CATEGORY LABEL	STRUCTURE MEASURE	S.E.	SCORE-TO-MEASURE AT CAT.	--ZONE--		50% CUM. PROBABLTY	COHERENCE M->C	C->M	ESTIM DISCR	
0	NONE		(1.83)	-INF	2.58		98%	98%		0
1	3.16	.46	3.35	2.58	3.82	2.87	87%	100%	1.42	1
2	4.15	.67	4.19	3.82	4.53	3.94	50%	33%	1.18	2
3	4.67	.71	4.91	4.53	5.48	4.52	25%	33%	1.79	3
4	4.38	.60	(6.04)	5.48	+INF	5.07	77%	70%	1.10	4

M->C = Does Measure imply Category?
C->M = Does Category imply Measure?

```
       CATEGORY PROBABILITIES: MODES - Structure measures at intersections
P      -+-------+-------+-------+-------+-------+-------+-------+-------+-
R  1.0 +000000000000000                                  444444444444444+
O      |                0000                          4444              |
B   .8 +                    00                      44                  +
A      |                      0                    4                    |
B   .6 +                       0                  4                     +
I   .5 +                        00               4                      +
L   .4 +                          11*11         4                       +
I      |                        11     02**    3                        |
T   .2 +                     11       2203**2333                        +
Y      |                1111     222 3**01122233333                     |
    .0 +**************************4       00****************************+
O      -+-------+-------+-------+-------+-------+-------+-------+-------+-
F      -8      -6      -4      -2       0       2       4       6       8
          PERSON [MINUS] ITEM MEASURE
```