MULTIPLE PERSPECTIVES OF THE FUNCTIONAL STATUS OF STROKE SURVIVORS AT 3 MONTHS POST-STROKE

by

Min-Mei Shih

BS, Chung Shang Medical University, 2003

MS, University of Pittsburgh, 2004

Submitted to the Graduate Faculty of

School of Health and Rehabilitation Sciences in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

University of Pittsburgh

2008

UNIVERSITY OF PITTSBURGH

SCHOOL OF HEALTHE AND REHABILITATION SCIENCES

This dissertation was presented

by

Min-Mei Shih

It was defended on

April 3, 2008

and approved by

James J. Irrgang, PhD, PT, ATC, Associate Professor, Department of Orthopaedic Surgery

Joan C. Rogers, PhD, OTR/L, Professor, Department of Occupational Therapy

Elizabeth R. Skidmore, PhD, OTR/L, Assistant Professor, Department of Occupational Therapy

Margo B. Holm, PhD, OTR/L, Professor, Department of Occupational Therapy
Dissertation Director

MULTIPLE PERSPECTIVES OF THE FUNCTIONAL STATUS OF STROKE SURVIVORS AT 3 MONTHS POST-STROKE

Min-Mei Shih, M.S.

University of Pittsburgh, 2008

Stroke is one of the leading causes of disability. Using an understandable measure to describe subsequent disabilities, namely, activities of daily living (ADL), is important for clinical practice. The three studies in this dissertation describe ADL task disability of stroke survivors at 3 months post-stroke, from multiple perspectives. The first study compared the constructs of five commonly used ADL measurement tools which used different scoring systems and assessment methods. Rasch analysis, using the partial credit model, confirmed that the performance-based and task-specific (criterion-referenced) ADL assessment, Performance Assessment of Self-Care Skills (PASS), had excellent unidimensionality for measuring independence in stroke survivors. It was also more valid and reliable than the other informant-based, and global non-summative (Glasgow Outcome Scale, 5-point [GOS5], Glasgow Outcome Scale, 5-point [GOS8], Modified Rankin Scale [mRS]) and global summative (Barthel Index [BI]) measures. The second study went on to develop an item difficulty hierarchy with the combined items from the PASS and the BI, and establish the person abilities of the stroke survivors. Rasch analysis and common person equating method revealed that the PASS was more difficult for the stroke survivors than the BI, and the participants had the greatest difficulty performing PASS instrumental ADL (IADL). The third study further delineated the independence of the stroke survivors with left and right hemispheric stroke (LHS and RHS) at the overall, domain, and task levels of the PASS. Rasch analysis, differential group functioning, and differential item functioning showed that the LHS

group performed significantly more independently than the RHS group in the functional mobility domain, and better, but not significantly better on the overall PASS, and the personal care, physical IADL, and cognitive IADL domains. The findings of clinically significant differences in specific tasks between the two stroke groups (side of lesion, gender, and age) will advance the knowledge related to specific disabilities of stroke survivors, especially for IADL tasks. Further studies were recommended to explore the independence of the stroke survivors in performing ADL subtasks, with more homogeneous samples and at multiple time points.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

I would like to thank many people who helped me to complete this dissertation. First, I would like to thank the chair of my dissertation committee, Dr. Margo B. Holm, PhD, OTR/L. Dr. Holm mentored me in multiple perspectives, including research studies, data analyses, teaching, and presentations, and guided me to overcome difficulties during the doctoral training and dissertation process.

I would also like to thank the members of my dissertation committee, Dr. Joan C. Rogers, PhD, OTR/L, for sharing research experiences and challenging my abilities to do research; Dr. Elizabeth R. Skidmore, PhD, OTR/L, for sharing clinical experiences and encouraging my interests in neurorehabilitation; and Dr. James J. Irrgang, PhD, PT, ATC, for expertise in Rasch analysis and judicious feedback. I also thank researchers with whom I have studied during my doctoral training for introducing their research to me.

Thank you to all of my peers and colleagues who offered assistance, encouragement, and comments during the dissertation process. I also thank my English tutors, Mr. Matthew Kessler, Mr. Willie Costello, and Ms. Susan Demo, for improving my English skills.

Finally, I would like to deeply thank my parents, Jin-Jia Shih and Hsueh-E Chou, brother, Min-Chi Shih, and my family members in Taiwan for being totally supportive throughout the dissertation process. I am humbled to share this accomplishment with them.

1.0 INTRODUCTION

Stroke (cerebrovascular accident or CVA) is a leading cause of disability in the United States. Approximately 700,000 people experience new or recurrent strokes each year. At 3-months post-stroke, 20% of stroke survivors require institutional care (American Heart Association, 2007). About 15% to 30% of stroke survivors are left with permanent disabilities. From the clinical point of view, it is important to clarify differences in disabilities associated with left and right hemispheric stroke (LHS and RHS). Rehabilitation practitioners are concerned not only with the level of impairment, such as limited motor or cognitive function, but also with the subsequent disabilities (Bernspang & Fisher, 1995).

Disabilities include limitations in the ability to independently perform daily activities (Carod-Artal, Gonzalez-Gutierrez, Herrero, Horan, & de Seijas, 2002), and a practitioner's ability to measure task disabilities is critical in clinical practice (Rogers, 1983). A functional status assessment that includes disability measures can enable practitioners to explain, confirm, or discriminate stroke survivors' performance in activities of daily living (ADL) (Rubenstein et al., 1988). In addition, because ADL is relatively objective, and relevant to patients, Duncan, Jorgensen, and Wade (2000) suggested that the primary functional status measure in stroke rehabilitation should focus on ADL. The most common ADL construct measured in rehabilitation is independence (Pamela W. Duncan, Jorgensen, & Wade, 2000; Rogers & Holm, 1998; van Boxel, Roest, Bergen, & Stam, 1995). The independence data from a functional status assessment can significantly influence the rehabilitation practitioner's understanding of a stroke survivor's disability and are often a key element in their clinical reasoning as they develop treatment plans (Dittmar, 1997; Rogers, 1983). In rehabilitation, having an understandable measure that both describes the functional status of a patient and guides intervention is extremely important (Smith & Taylor, 2004).

The current study was done to advance knowledge about ADL disability measures, and ADL disability in stroke survivors, at 3 months post-stroke. As a secondary outcome, a hierarchy of ADL tasks among stroke survivors was developed to provide guidance for stroke rehabilitation. The general aims of this study were to:

(1) confirm, correlate, and compare the measurement properties of five functional status instruments that are used often with stroke survivors, and use different scoring systems to rate ADL disability,

(2) compare specific disability items for stroke survivors at 3-months post-stroke, and

(3) describe the differences in disabilities between persons with LHS and RHS and compare the functional impact of LHS and RHS on specific ADL tasks.

Chapters 2, 3, and 4 present research studies focusing on ADL functional status measures and the ADL independence of stroke survivors. In Chapter 2, the measurement constructs of the five measures (the Glasgow Outcome Scale, 5-point version [GOS5], the Glasgow Outcome Scale, 8-point version [GOS8], the Modified Rankin Scale [mRS], the Barthel Index [BI], and the Performance Assessment of Self-Care Skills [PASS]) were confirmed for each measure separately. Then the scores from each measure were correlated to understand the relationship among the five measures. Finally, to compare item difficulty among the five measures for independence, the scores from five measures were converted to a single metric, using Rasch analysis. Then person ability and item difficulty were mapped for the combined measures to illustrate the interaction between item difficulty and person ability.

Even though they had been partially discussed in the previous chapter, in Chapter 3, we extracted out the global non-summative measures (GOS5, GOS8, and mRS), to compare the difficulty of the items in the rehabilitation-relevant BI and PASS tools in depth. This is because the BI is frequently considered the gold standard even when it is informant-based. To do this, we again combined the BI and PASS items into a single metric, using Rasch analysis. Differences in the difficulty of common items in these two assessments were also examined.

Chapter 4 begins with a systematic review of the literature comparing differences in ADL independence between stroke survivors with LHS and RHS. The second part of this chapter examines the differences in overall ADL independence, ADL domain independence, and specific ADL task independence between stroke survivors with LHS and RHS on the PASS. The results from the current study are compared with findings from the systematic review and other relevant literature, and clinically implications of the findings are reviewed. In Chapter 5, the results of the three studies are summarized. The implications of our findings for ADL performance of stroke survivors at 3 months post-stroke are also delineated.

2.0 MEASURING STROKE SURVIVORS' ACTIVITIES OF DAILY LIVING: FIVE PERSPECTIVES

2.1 BACKGOUND AND SIGNIFICANCE

Stroke (cerebrovascular accident or CVA) is one of the leading causes of disability in the United States, with approximately 700,000 people experiencing new or recurrent strokes each year. The American Heart Association (2007) reports that 20% of stroke survivors require institutional care at 3 months post-stroke and 26% of the ischemic stroke survivors are still institutionalized in a nursing home at 3 months post-stroke. About 15% to 30% of stroke survivors are left with permanent disabilities (American Heart Association, 2007). The disabilities related to stroke include functional limitations in the "capacity to carry out any activity within the normal range achievable by a human being," (Carod-Artal, Gonalez-Gutierrez, Herrero, Horan, & de Seijas, 2002, p. 207). Clinically, a functional status assessment is used to identify disabilities in activities of daily living (ADL) such as eating, dressing, and managing financial tasks. Functional status assessments allow rehabilitation practitioners to document patients' performance of everyday activities to "help to explain, confirm, or cast doubt" (Rubenstein, Calkins, Greenfield, et al., 1988, p. 563) on their functional status. However, multiple functional status instruments are used in rehabilitation clinics and their scoring systems often vary. The present study aimed to confirm, correlate, and compare the measurement properties of five functional status instruments that are used often with stroke survivors, and use different scoring systems to rate ADL disability. Participants were stroke survivors at 3 months post-stroke.

2.1.1 Stroke survivors and functional status

Functional status refers to a person's ability to perform ADL and fulfill social roles, at a specific point in time (Rubenstein et al., 1988). Determining what to measure when assessing "functional status" can be somewhat ambiguous. One common approach used in rehabilitation is to evaluate independence in performing functional activities, namely, ADL (Pamela W. Duncan et al., 2000; Rogers & Holm, 1998; van Boxel et al., 1995). Basic ADL (BADL) and instrumental ADL (IADL) are the two fundamental categories of ADL (Moskowitz & McCann, 1957; Reuben & Solomon, 1989). BADL refers to self-maintenance activities involving functional mobility and personal care, such as ambulation or wheelchair mobility, transfers, feeding, hygiene, toileting, and bathing (Katz, Ford, Moskowitz, Jackson, & Jaffe, 1963; Rogers & Holm, 1998). IADL refers to more complex activities involving home management tasks and tasks required for independent living in the community, such as shopping, cooking, housekeeping, laundry, use of transportation, managing money, managing medicine, and use of the telephone (Lawton & Brody, 1969; Rogers & Holm, 1998). Functional status data provide a means for rehabilitation practitioners to communicate, document progress, and compare or monitor the functional status of stroke survivors (Portney & Watkins, 2000; Rubenstein et al., 1988). Functional status data significantly influence the rehabilitation practitioner's understanding of a stroke survivor's disability and are often a key element in their clinical reasoning as they develop treatment plans (Dittmar, 1997; Rogers, 1983). The data provide the most concrete and detailed information about the disabilities of stroke survivors (van Boxel et al., 1995). Duncan, Jorgensen, and Wade

(2000) also suggested that ADL should be the primary functional status measure in stroke rehabilitation due to their relative objectivity, simplicity, and relevance to patients. Moreover, in rehabilitation, it is extremely important to have an understandable measure that both describes the functional status of a patient and guides intervention (Smith & Taylor, 2004).

2.1.2 Functional status scoring systems (4 types)

Functional status scoring systems can be global or task specific. They can be divided further into four types: (1) global non-summative scores, (2) global summative scores, (3) global average scores, and (4) task-specific scores (see Table 2-1). Global non-summative single scale scoring is usually an ordinal scale and reflects increasing or decreasing levels of overall functional status. Examples of instruments using a non-summative single scale scoring are the modified Rankin Scale (mRS) and the Glasgow Outcome Scale (GOS). Global summative scoring yields a summary score by simply adding up the obtained ratings on all items to indicate the severity of disability. Examples of instruments using global summative scoring are the Barthel Index (BI) and the Functional Independence Measure (FIM) (Ravaud, Delcey, & Yelnik, 1999; Ring, Feder, Schwartz, & Samuels, 1997; van Hartingsveld, Lucas, Kwakkel, & Lindeboom, 2006). The third type of global scoring system uses an average of items (Holm & Rogers, 1999). An example of an instrument using item and domain (e.g., BADL and IADL) grand means is the Performance Assessment of Self-Care Skills (PASS), in which a total grand mean is obtained by averaging the scores of the tasks (Holm & Rogers, 1999).

Table 2-1: Examples of Global Functional Status (non-summative, summative and global average) Scoring Systems

Functional Status Measures	Type	Sample Scoring	
		Score	Description
Modified Rankin Scale (mRS)	Global Non-summative	1	No significant disability despite symptoms: able to carry out all usual duties and activities
		5	Severe disability: bedridden, incontinent, and requiring constant nursing care and attention
Glasgow Outcome Scale (5 pt. version) (GOS5)	Global Non-summative	3	Severe disability
		5	Good recovery
Glasgow Outcome Scale (8 pt. version) (GOS8)	Global Non-summative	2	Non-sentient, not obeying commands, no verbal response, no meaningful response: May have sleep-wake rhythm, may have spontaneous eye opening and ability to follow moving objects, may swallow food
		7	Capable of resuming normal occupational and social activities; there are minor physical or mental deficits or complaints

Table 2-1 (Continued)

<table>
<tr><td rowspan="3">Barthel Index (BI) Items
(20 pt. version)

Bowels, Bladder, Grooming, Toilet use, Feeding, Transfer (bed-chair), Walking, Dressing, Stairs, Bathing</td><td rowspan="3">Global Summative</td><td colspan="2">Higher scores indicate greater independence in activities of daily living. Scores are derived by adding up item scores.
Sample item: BI10 - Bathing</td></tr>
<tr><td>1</td><td>Independent (May use bathtub, shower or sponge bath; Subject must be able to perform all functions without another person being present</td></tr>
<tr><td>0</td><td>Dependent (Cannot meet criteria)</td></tr>
<tr><td rowspan="3">Performance Assessment of Self-Care Skills (PASS) Task Items

Bed transfers, Stair use, Toilet transfers, Oral hygiene, Bathtub/Shower transfers, Trimming toenails, Dressing, Shopping, Bill paying, Checkbook balancing, Mailing, Carrying garbage, Telephone use, Medication management, Changing bed linens, Obtaining information-auditory, Obtaining information-visual, Small repairs, Sweeping, Indoor walking, Home safety, Playing bingo, Stovetop use, Using sharp utensils, Cleanup after meal preparation</td><td rowspan="3">Global Average [Task-specific]</td><td colspan="2">Higher scores indicate greater independence in activities of daily living. Global average scores are derived by calculating a grand mean from item means.</td></tr>
<tr><td>3</td><td>Independent (No assists given for task initiation, continuation, or completion)</td></tr>
<tr><td>0</td><td>Dependent (Total assistance given, or continuous physical guidance or physical support during tasks)</td></tr>
</table>

Unlike the three global scoring systems, a task specific scoring system uses the ratings of single tasks which have critical and essential criteria (i.e., a criterion-referenced test), and each item can stand alone. An example of a task-specific instrument is the PASS. In the PASS, each task contains critical subtasks that have criteria for rating ADL performance. For the independence scale of the PASS, the rating is based on the level of assistance provided for each subtask (Holm & Rogers, 1999). For example, the PASS item "Stovetop Use" consists of 11 subtasks, with clear, observable criteria to be rated (e.g., Turns burner off promptly [+/-- 1 minute of removing soup from burner]) (see Table 2-2).

Table 2-2: Example of Performance Assessment of Self-Care Skills (PASS) Item Rating Criteria

Sample task: Stovetop use	
1	Open soup can correctly (cut is even, entire top is off or < ½ is retained in one place)
2	Removes/handles soup can lid correctly (lifts lid with knife; punches lid into can; does not cut finger)
3	Pours/spoons soup into pan without spilling (no soup on patient, counter, or floor)
4	Adds water correctly (add 1 can of water; does not spill on self, floor)
5	Places pan on correct stove burner (burner closest to pan size)
6	Turns burner on correctly (manipulates knob for burner that soup is on or is placed onto later; sets control on medium to high)
7	Monitors soup adequately (stirs; alters heat as necessary, soup does not stick on pan, checks to make sure soup temperature is hot rather than lukewarm to touch or taste or that soup boils/bubbles)
8	Removes pan from burner when soup is still hot (steam can be seen rising from pan; checks to make sure soup temperature is hot rather than lukewarm to touch or taste)
9	Turns burner off promptly (+/- 1 minute of removing soup from burner)
10	Transports & pours soup into bowls correctly (uses mitt under pan or slides pan across counter for stability if weakness or tremor present; does not spill on floor; only minor drips on counter)
11	Transports bowls to table correctly (uses mitt under bowl or uses cart if weakness, tremor or instability present; uses bowl rim to carry; does not spill on floor)

2.1.3 Functional status scoring systems: Advantages and disadvantages

There are advantages and disadvantages to the clinical utility of the four types of scoring systems. The commonly used instruments that use a non-summative single scale scoring system in stroke studies are good for global estimates in population based studies, but often combine different concepts of stroke disability. For instance, level 0 of the mRS is "No symptoms at all" whereas a 4 on the mRS refers to "moderate to severe disability," thus combining stroke symptoms and disability. The GOS also mixes concepts such as "vegetative state" (score of 2) and "moderate disability" (score of 4). The utility of such scores for rehabilitation practitioners is thus limited (Kasner, 2006; Lindeboom, Vermeulen, Holman, & Haan, 2003).

Likewise, with data from summative scoring systems, rehabilitation practitioners can only obtain a broad impression of the patient's performance (i.e., a higher sum score on the BI indicates "more independence in BADL"). A limitation of summative scoring systems is that the exact functional ability of patients on specific tasks cannot be determined unless go back to items, thus decreasing the instrument's usefulness. For example, two patients can have a BI score of 70, with one patient unable to walk, dress, and bathe independently, and the other unable to feed, or be continent of bowel and bladder. Overall, although both non-summative single scale and summative scoring systems explain functional status globally, they do not provide enough detail to guide intervention. The consequence of no specificity for ADL task performance is that there is not enough information to guide the clinical reasoning of rehabilitation practitioners for developing suitable interventions (Kasner, 2006).

Compared to the two other global scoring systems, averages can provide more specific information about disability if averages are derived from a task-specific scoring system with clear criteria for rating performance (Rogers et al., 2003). Although still global in nature, averages can provide greater direction to rehabilitation practitioners, especially if grand means are derived from subtasks of a single task (stovetop cooking), or from like tasks (e.g., BADL domain). Task-specific scoring also results in better reliability between raters, especially when criteria are stated clearly and are observable (Rogers, Holm, Beach, Schulz, & Starz, 2001). It is unclear if a global averaging scoring system for all items of a criterion-referenced instrument can provide better functional status data than global non-summative or global summative scoring systems.

2.1.4 Functional status scales

In addition to the issue of scoring systems, most ADL measurements utilize ordinal scales. Linedeboom, Vermeulen, Holman, and Haan (2003) mentioned that traditional statistical methods that assume the data are on an interval scale cannot solve the difficulties related to ordinal scale scores for ADL assessments. Owing to the ordinal structure of the item set, the scores at the same scale levels may actually be different because the distances between items along the scales are not equal and because they may represent varied meanings (Merbitz, Morris, & Grip, 1989). Furthermore, when classical statistical methods are applied to such scales, "examinee characteristics and test characteristics cannot be separated" (Hambleton, Swaminathan, & Rogers, 1991, p. 2). This conceptual problem influences the interpretation of results and leads to difficulty when practitioners try to compare performance among different tests. Rasch analysis was developed to address such limitations. In simplistic terms, Rasch analysis uses logarithmic transformation to convert ordinal data into interval data, yielding an output called a logit (log odds unit) scale (Bond & Fox, 2007; Hambleton, Swaminathan, & Rogers, 1991; Tesio, 2003; Wright & Linacre, 1989). Rasch analysis is advantageous for interpreting functional status, because it can "use a single synthetic index to describe the full spectrum of functioning" (Barberger-Gateau, Rainville, Letenneur, & Dartigues, 2000, p. 310). Rasch analysis also yields two parameters – item difficulty and person ability – that allow item and person information to be separated and interpreted at the same time using the same measurement units (Hays, Morales, & Reise, 2000; Tesio, 2003).

2.1.5 Comparing functional status instruments used with stroke survivors

Not only can Rasch analysis be used to convert ordinal scales to interval scales, the Rasch family of models provides an opportunity to compare functional status instruments that use different scoring and scaling systems. To use Rasch analysis, the assumption of unidimensionality in the Rasch model requiring the measurement to "focus on one attribute or dimension at a time" (Bond & Fox, 2007, p. 32) should be met. However, there has been criticism when groups have been compared using item difficulty estimates (Wright, 1977). That is, a meaningful and useful measurement exists only when it measures one attribute (e.g., independence in daily tasks). More than one attribute would confuse the prediction or implication of obtained scores (Bond & Fox, 2007). Also, item difficulty or the item hierarchy should not change depending on person ability and should be stable across groups (Dallmeijer et al., 2005; Tesio, 2003). Therefore, to compare performance among groups, it is appropriate to compare groups by using person ability estimates. To date, no evidence could be found that compared different scoring systems used to measure the functional status of stroke survivors, nor were any studies found that compared them on person ability estimates from Rasch analysis.

2.1.6 Aims and Hypotheses

By analyzing a wide range of ADL performance (BADL and IADL), the present study aimed to confirm the measurement properties of each of the five functional status instruments, and then correlate the obtained scores. Based on more advantages for the scoring systems of the BI and the PASS than those of the global non-summative measures (the GOS5, GOS8, and mRS), we hypothesized that the relationship between the PASS and the BI scores would be the strongest.

That is, there would be a stronger relationship between the PASS and the BI scores than between the PASS and the GOS5, GOS8, and mRS or between the BI and the GOS5, GOS8, and mRS. Finally, the present study aimed to compare the data from global and task-specific functional status measures with different scoring systems after we combined the data from the five tools into a single metric, using Rasch analysis.

2.1.7 Participants

Data for the current study were from a large prospective stroke outcome study: The Use of Radiological Data to Describe, Differentiate, and Predict Impairment, Disability, and Quality of Life in Stroke/ Transient Ischemic Attack Survivors (IRB# 010593) The inclusion criteria were: (a) admission to the University of Pittsburgh Medical Center (UPMC) Presbyterian Hospital (b) diagnosis of acute stroke or a transient ischemic attach (c) radiological data available from admission assessments (e.g., CT scan, MRI, and or Xenon quantitative cerebral blood flow [ExCT qCBF]) or pharmacological interventions (tPA), and (d) physician approval. There were no exclusion criteria based upon gender, race, ethnicity, or HIV status. However, children were excluded from the study because a pharmacological intervention, tissue plasminogen activator (tPA), used in the larger study was not approved for use in children. For the current study, only data from the 3-months assessment, for subjects with diagnoses of ischemic or hemorrhage stroke and deficits in the right or left hemisphere, were included.

2.1.8 Procedures

The recruitment procedures in the large prospective stroke outcome study were approved by University of Pittsburgh Institutional Review Board and followed the guidelines of confidentiality under HIPAA. The study team requested informed consent directly from the patient or the appropriate proxy within 24 hours after admission to UPMC Presbyterian hospital with a primary diagnosis of acute stroke. After eligibility requirements were verified, demographic data were gathered by the attending physician or nursing staff. The data were documented in the medical records. At 24 hours, 5 days, and 3, 6, 9, and 12 months post acute stroke event, the ADL assessments were administered by the staff of the Department of Occupational Therapy, who were trained by the Principal Investigator (Holm) in the assessment procedures to a minimum interobserver standard of >90% with the criterion assessor (IRB#010593). The 3-months ADL assessments took place where the person was residing at the time, within a 150-mile radius of Pittsburgh**.**

2.1.9 Instruments

ADL functional status of participants 3 months post-stroke was measured with global and criterion-referenced task specific ADL measures. The global ADL measures were two versions of the Glasgow Outcome Scale (GOS) (Jennett & Bond, 1975), the modified Rankin Scale (mRS) (Rankin, 1957; van Swieten, Koudstaal, Visser, Schouten, & van Gijn, 1988), and the Barthel Index (BI) (Wade & Collin, 1988). The global average and task specific ADL measure was the Performance Assessment of Self-Care Skills (PASS) (Holm & Rogers, 1999).

2.1.9.1 Glasgow Outcome Scale, 5-point version (GOS5)

The GOS was used to rate communication and independence in activities of daily living. It is a commonly used outcome classification scale for disability and has demonstrated adequate reliability and validity (Jennett, Snoek, Bond, & Brooks, 1981). The original GOS is a 5-point scale, scores range from 1 (death) to 5 (good recovery). Between the scores 1 and 5, the score 2 refers to vegetative state, 3 refers to severe disability, and 4 refers to moderate disability.

2.1.9.2 Glasgow Outcome Scale, 8-point version (GOS8)

The original GOS was extended to an 8-point scale, where each of the scores 3, 4, and 5 in the GOS5 expands to two categories. For example, the category "moderate disability" (score 4) in the GOS5 expands to two categories in the GOS8, "independent in activities of daily life, for instance can travel by public transport, not able to resume previous activities either at work or socially; despite evident post-traumatic signs, resumption of activities at a lower level is often possible" (score 5) and "post-traumatic signs are present, however, resumption of most former activities either full-time or part-time is possible" (score 6). The extended scale is more sensitive than the GOS5 for measuring various levels of physical and mental disabilities. It has also demonstrated adequate reliability and validity {Jennett, 1981 #174. The extended GOS scores range from 1 (death) to 8 (good recovery).

2.1.9.1 Modified Rankin Scale (mRS)

The mRS is a one-item, 5-point scale used to rate disability and need for assistance. It is the most commonly used outcome classification scale for disabilities and handicaps after stroke and has demonstrated adequate interobserver reliability ($\kappa = .65$, $\kappa_W > .91$) (van Swieten et al., 1988; C. D. A. Wolfe, Taub, Woodrow, & Burney, 1991). The mRS has excellent agreement with the

GOS5 (94%) and BI (87%), which confirmed its construct validity (Tilley, Marler, Geller, & National Institute of Neurological Disorders and Stroke [NINDS] rt-PA Stroke Trial Study, 1996). The mRS was rated from data gleaned from the other measures. Rating scale scores range from 0 (no symptoms at all) to 5 (severe disability).

2.1.9.2 Barthel Index (BI)

The BI was used to rate independence in 10 basic activities of daily living (BADL) (bowel and bladder control, toileting, transfers, wheelchair mobility, stairs, grooming, bathing, feeding, and dressing). It has been used extensively in clinical research, with high interrater reliability ($r_s > .89$, $p < .001$) and correlates highly with performance-based functional measures ($r_s > .80$, $p < .01$) (Shinar, Gross, Bronstein, & Licata-Gehr, 1987). Factor analysis has confirmed that the BI has construct validity, indicating that the BI items measure the same domain. The BI was also shown to measure similar ADL as other ADL measures, such as Katz Index of ADL (concurrent validity) (Wade & Collin, 1988). In addition, the stroke survivors who had higher BI scores at discharge had better outcomes (predictive validity) (Granger, Hamilton, & Graesham, 1988). The BI can be administered in several formats, including self-report interview, chart review and clinical observation. Each item has its own rating criteria, which yield different scores, ranging from 0 to 3. Lower scores indicate less independence during the activity, while higher scores indicate more independence.

2.1.9.3 Performance Assessment of Self-Care Skills (PASS)

The PASS is a criterion-referenced, performance-based instrument, which includes 26 tasks and 163 subtasks. The 26 tasks are further categorized into 4 domains, functional mobility (FM) (5 items, 28 subtasks), personal self-care (PC) (3 items, 26 subtasks), physical-instrumental activities of daily living (PIADL) (4 items, 22 subtasks), and cognitive-instrumental activities of daily living (CIADL) (14 items, 87 subtasks) (see Table 2-3). The examiner rates the ADL performance of the participants based on established criteria for each task, or its subtasks. Examiner verbal instructions and placement of task objects are standardized. The independence score for each subtask ranges from 0 (unable to perform task independently) to 3 (independent) (see Table 2-4), which are based on the frequency (e.g., occasional or continuous assists) and level (e.g., verbal or physical) of assistance provided by the examiner. Assistance is provided only when needed, starting from the least assistive prompt to the most assistive and intrusive prompt (see Table 2-5). A mean of the subtask independence scores yields the independence score for each task. To represent categories of each PASS task for the Rasch analysis, integer scores were yielded from recoding the averaging independence scores (i.e., The averaging scores from 0.01 to 1.50 were recoded into 1, and from 1.51 to 2.99 were recoded into 2, whereas the averaging scores of 0 and 3 remained the same for the following Rasch analysis). Because the data for the oven-use task was not always available, only 25 tasks from the PASS were included in the analyses. Although ratings of safety (i.e., the personal or environmental risks when performing a task) and adequacy (i.e., the level of efficiency of task initiation, continuation, and completion, and the degree of match between the end product and criteria identified as acceptable quality when performing a task) were rated with the PASS, the current study was delimited to the construct of independence, and only those PASS data were analyzed.

Validity of the PASS was referenced to common geriatric ADL/IADL instruments, the Lawton and Brody's (1969) Scales for Instrumental Activities of Daily Living, the Older Adults Resources and Services (OARS) ADL Scale {Fillenbaum, 1988 #175}, the Comprehensive Assessment and Referral Evaluation (CARE) (Gurland et al., 1977), and the Functional

Assessment Questionnaire (Pfeffer, Kurosaki, Karrah, Chance, & Filos, 1982). Test-retest reliability on two consecutive days was r = 0.96. Unidimensionality of the PASS independence construct was investigated using exploratory factor analysis. Factor analysis using SPSS 12.0 examined independence scores for 26 tasks of the PASS for 1158 subjects, including populations with depression, osteoarthritis, cardiopulmonary disease, and dementia, macular degeneration, and stroke as well as a cohort of well-elderly. The largest eigenvalue for the 26 tasks, accounted for over 37% of the variance, and was 3.44 times larger than the second largest eigenvalue. Because the investigation of the PASS independence data revealed the presence of one dominant construct, the unidimensionality of the PASS independence data was established (Chisholm, 2005).

Table 2-3: Performance Assessment of Self-care Skills (PASS) Tasks, by Domain

Domain / Task
Functional Mobility (FM)
Bed transfers (move from prone to supine position and rise from bed)
Stair use (ascend and descend stairs)
Toilet transfers (sit and rise from a toilet)
Bathtub/shower transfers (enter and exit tub and/or shower)
Indoor walking (walk indoors)
Personal Self-care (PC)
Oral hygiene (clean teeth, dentures and/or mouth)
Trimming toenails (groom toenails)
Dressing (don and doff upper body and lower body clothing)
Physical Instrumental Activities of Daily Living (PIADL)
Carry garbage (bend, lift and carry garbage sack)
Changing bed linen (put on bed linens)
Sweeping (clean spillage on the floor using a broom and a dust pan)
Cleanup after meal preparation (perform clean up tasks after meal preparation)
Cognitive Instrumental Activities of Daily Living (CIADL)
Shopping (select and purchase grocery items)
Bill paying (write checks for sample utility bills)
Checkbook balancing (balance a checkbook after writing checks)
Mailing (prepare envelopes for mailing checks)
Telephone use (use telephone to obtain information)
Medication management (read medication information and organize medication according to prescription)
Obtaining information: auditory (obtain information from a radio announcement)
Obtaining information: visual (obtain information from a newspaper)
Small repairs (repair a flashlight)
Home safety (identify and correct hazards or problems in home safety situations)
Playing bingo (play bingo)
Oven use (cook muffins in an oven)
Stovetop use (cook soup on a stovetop)
Using sharp utensils (cut an apple with a sharp knife)

Note. This table (Chisholm, 2005, p. 10) is used with permission.

Table 2-4: Performance Assessment of Self-care Skills (PASS) Independence Scoring Criteria

SCORE	CRITERIA
INDEPENDENT PERFORMANCE	
3	No assists given for task initiation, continuation, or completion
2	No Level 7-9 assists given, but occasional Level 1-6 assists given
1	No Level 9 assists given; occasional Level 7 or 8 assists given, or continuous Level 1-6 assists given
0	Level 9 assists given, or continuous Level 7 or 8 assists given; or unable to initiate, continue, or complete subtasks or task
DEPENDENT PERFORMANCE	

Note. This table (Chisholm, 2005, p. 11) is used with permission.

Table 2-5: Performance Assessment of Self-care Skills (PASS) Prompt Hierarchy

	LEVEL	PROMPT	DESCRIPTION
LEAST RESTRICTIVE			
	0	No assistance of any type	Person initiates, continues, completes subtask without assistance
VERBAL	1	Verbal support	Encouragement
	2	Verbal non-directive	Cue to alert that something is not right
	3	Verbal directive	Tell person what to do next
GESTURE	4	Gestures	Point at task object
	5	Task/environmental rearrangement	Break task down into manageable components
	6	Demonstration	Assessor demonstrates/person follows
PHYSICAL	7	Physical guidance	"Hands down" – move body part into place
	8	Physical support	"Hands up" – lift body part/clothes/support
	9	Total assist	Assessor does task or subtasks for the person
MOST RESTRICTIVE			

Note. This table (Chisholm, 2005, p. 12) is used with permission.

2.2 DATA ANALYSIS

2.2.1 Data preparation

Descriptive statistics were calculated for all demographic variables. Participants' demographics, as well as pathologies or impairments related to stroke, were analyzed using SPSS version 14.0. Missing data were handled using linear interpolation, an SPSS estimation method for replacing missing values. Before replacing missing values, the dataset was divided into two sets according to like subjects. That is, data included only participants with left hemispheric stroke or with right hemispheric stroke. The last valid value before the missing value and the first valid value after the missing value were used for the interpolation. The missing value is the median value of its valid, surrounding values. After the missing values were replaced from like subjects in appropriate datasets, they were merged (SPSS Inc., 2005). This procedure was done before running Rasch analysis in WINSTEPS version 3.64.1.

2.2.2 Category analysis of global non-summative measures

The three global non-summative measures (GOS5, GOS8, and mRS) were composed of single scales. Their raw scores could not be analyzed using Rasch analysis because the probabilities of the item responses and person abilities would be omitted (Bond & Fox, 2007). Therefore, category analyses for these measurements were performed. The frequency and the percentage of the observed responses of each category in these measurements were calculated. The results were used for comparison with the responses of the BI and the PASS.

2.2.3 Rasch analysis

2.2.3.1 Logits of item difficulty and person ability

Rasch analysis was used to transform ordinal data, from the PASS and BI, into interval data, or logits (log odds units). This enabled the unit intervals between locations on the item logit scale to have a consistent value or meaning, and therefore the transformations allowed for comparison of the relative difficulty of tasks to other tasks or the relative ability of persons to other persons. The item difficulty logit of an item in a measurement tool represents its underlying difficulty calculated from the total number of people in the sample who pass the item successfully. The person ability logit of a person in a sample represents underlying ability, or the performance on the measurement tool, calculated from the total number of items in a measurement tool which the person passes successfully. "A person's ability is [also] defined as the probability of that person having a 50 % chance of getting an item of given difficulty right" (Wolfe et al., 2000, p. 1992).

2.2.3.2 Hierarchies of item difficulty and person ability

Because Rasch analysis allowed all testing items to be placed on one scale, a hierarchy of easiest to most difficult tasks was created. Likewise, all persons could also be placed on one scale, thus creating a hierarchy of persons based on best to worst performance, or the likelihood of performing or not performing tasks independently. The current study used a score of zero as the midpoint of difficulty. For the scores from the PASS, the items with the more positive logit values were harder to perform while those with a more negative value were easier. In contrast, persons with more positive logit values had a greater likelihood of performing tasks independently than persons with lower or negative logit values. The BI independence scores were also entered into Rasch analysis. Due to the different scaling of each item of the BI, the Partial Credit Rasch Model (PCM) was applied. The PCM allows items with various rating scales to be analyzed and keeps the distances between rating scales constant, whereas the regular Rasch model can only be applied for items with the same rating scale. solves the different intermediate levels that come from different numbers of responses for different items on the same instrument (Bond & Fox, 2007; Linacre, 2007). Similar to the interpretations of the estimates from the PASS scores, the items of the BI with the more positive logit values were harder to perform while those with a more negative value were easier. Likewise, the persons with more positive logit values had a greater likelihood of performing tasks independently than persons with lower or negative logit values. The PCM in Rasch analysis was used to analyze scores of the five ADL measures, the GOS5, GOS8, mRS, BI, and PASS (5-ADL-measures). The purpose of performing the analyses was to compare the items in these measures, especially to compare the constructs of the global non-summative measures (the GOS5, GOS8, and mRS) to those of the BI and the PASS. Therefore, these items were combined and treated as a new evaluation tool, in order to examine item difficulty and person ability across measures. For checking the scales of the global non-summative measures (i.e., GOS5, GOS8, and mRS), the scores were recoded to yield consistent scales. That is, every scale of these measures was transformed to begin with 0, the same as the scales of the PASS and BI. This procedure prevented the Winsteps program from ignoring data, or treating some data as not administered. The mRS scale was also reversed before the analysis to be consistent with the direction of other scales (i.e., score 0 indicates the best performance in mRS). Therefore, the logits of the item difficulty and the person ability of the 5-ADL-measures can be interpreted similar to those of the PASS and the BI.

2.2.3.3 Diagnosis (construct validity): Fit statistics

There are two primary assumptions in the Rasch model, derived from item response theory (IRT): unidimensionality (i.e., focusing on one attribute or dimension at a time) and local independence (i.e., no significant relationship between the test items, although the latent trait of the examinees is homogeneous) (Hambleton et al., 1991; Scherbaum Jr., 2003). Hambleton et al. (1991) stated that the assumption of local independence is automatically met when the assumption of unidimensionality is met. Because Rasch analysis is a model-based approach, fit statistics are performed to detect discrepancies between the Rasch model and the data collected in practice. The fit statistics provide information about how to interpret the data precisely if any discrepancy occurs. The estimates of item difficulty or person ability are along a logit scale, and "each of these estimates has a degree of error associated with it" (Bond & Fox, 2007, p. 41). If the data fit perfectly with the linear logit scale, a unidimensional interval measurement scale, the tool to test human performance is considered stable and reliable. The discrepancy between the Rasch model and the data is reported by the infit and outfit statistics. Both infit and outfit statistics are reported as mean squares (mean of squared residuals) and represent the variation between observed responses and model-predicted responses. The residuals are the differences between the model expected value and the actual performance for an item. The value of mean squares, in the form of a chi-square format, is composed of an expected value of +1 and a range from 0 to positive infinity. For instance, an infit mean square error value of 1.30 indicates 30% more variation in the observed data than the Rasch model predicted. For clinical observation tools, such as PASS, a reasonable task mean square error range for infit and outfit statistics is 0.5 to 1.7.

Misfit tasks with both infit and outfit mean square error values greater than the suggested level, ≤ 0.5 or ≥ 1.7, are required to be examined more closely. The infit statistic "gives relatively more weight to the performance of persons closer to the item value" (p. 57). In contrast, the outfit statistic is not weighted and so is more sensitive to the influence of outlying scores. Thus, more attention is usually paid to the infit statistics than the outfit statistics (Bond & Fox, 2007).

2.2.3.4 Diagnosis (construct validity): Principal component analysis of Rasch residuals

Principal component analysis (PCA) of Rasch residuals is an advanced method of determining the deviations from the assumption of unidimensionality in the Rasch model. Potential multi-dimensions of the measurement construct, other than the main dimension (e.g., measuring ADL functions in this study), would be revealed by using PCA (Bond & Fox, 2007). The PCA yields empirical values and modeled values that represent the variances explained or unexplained by the model. Among the unexplained variances, up to 5 contrasts (or additional dimensions) are identified (Linacre, 2007). An elbow-shaped line is plotted according to the variance log-scale, the percentage for each variance or contrast, against the variance components. After extracting the explained variance, PCA provides a unique Rasch factor analysis. Therefore, in addition to the percentages for these explained or unexplained variances, the eigenvalues are calculated to represent the amount of variance explained by each contrast. Because each item of the measure is independent and shares one unit of randomness in the analysis (i.e., eigenvalue), the strength of a contrast can be interpreted as the number of items out of the total items in the measure. The PCA limits further analysis to the first contrast, which is the largest secondary dimension of the measures. The first contrast identifies a common variance of subset items in the residuals. This Rasch analysis principal component analysis concept is different from traditional factor analysis.

The factor loading of each item distinguishes the common variance of subset items. The analysis also identifies the contrast deviations in positive loadings against negative loadings for the items in the residuals. A plot of residual factor loadings against the item difficulty logits is generated. The clustered items in each half of the plot illustrate items with potential common latent characteristics. This plot helps to identify whether the clustered items group at a particular difficulty level. The latent characteristics in the residual plot need to be interpreted based on their commonalities and that of the measurement tool and its construct. Moreover, the items with larger variances that are unexplained in the primary Rasch analysis are located at the top of the map, indicating higher factor loadings for these items. Further investigations are also recommended for clustered items with factor loadings greater than zero. The ratio of the eigenvalues of a contrast to the units (in eigenvalues) of the explained variance in the Rasch model, namely the factor sensitivity ratio, describes the impact of the item residuals in this contrast (a subscale) to the stability of the unexplained variances after the primary Rasch measure was extracted. The factor loadings and the factor sensitivity ratio together provide empirical evidence about the influence of the minor trait (Bond & Fox, 2007; Linacre, 2007). There are rules of thumb to discriminate whether the whole set of data is good and represents the unidimensionality of the total measurement constructs, or whether the residuals are only considered as *noise*. These rules consist of the following: (1) the explained variance of the model is greater than 60%, (2) that the unexplained variance of the first contrast is less than 5% or (3) the eigenvalue of the unexplained variance by the first contrast is smaller than 3.0 (Linacre, 2007).

2.2.3.5 Diagnosis (construct validity): Category function

Because the rating scale affects the quality of measuring and further influences the interpretation of results, checking the category function is essential. The category refers to the responses to an item in a measurement tool. For example, the PASS has 4 independence rating categories for a task (see Table 2-4). There are several characteristics that identify a well-functioning category: category frequency, average measures, thresholds, and category fits.

Category frequency is the percentage of observed ratings in the measurement tool for a sample. A category with low frequency is the result of fewer observations for this category, which indicates this category may be an unnecessary or redundant category. The minimal number of responses per category is recommended to be 10.

Average measures are the average of the person ability logits for people in the sample who have responses to a particular category, across all tasks. It is expected that the sizes of the average measures increase as the category value increases. That is, on average, people with higher abilities are represented in the higher category, and vice-versa.

Thresholds are "the difficulties estimated for choosing one response category over another" (Bond & Fox, 2007, p. 223-224) (i.e., how difficult it is to observe "dependent," score 0, over "independent," score 3). It is expected that the values of the thresholds increase with category values. The probability curves are also plotted to illustrate if these categories work satisfactorily for a sample. Each category is represented by a probability curve, and the values of the thresholds are the points at which adjacent category curves intersect. In the graph, each curve should have a distinct peak to demonstrate that it is the most probable category for a certain portion of the measured variable. If there is a flat curve, but it spans a large portion of the variable, this category is still useful. If, however, the curve is flat and over-shadowed by another curve, the category may be redundant. In addition, the threshold distance, the magnitudes of the distances between adjacent threshold estimates, should be neither too short nor too long. The recommended magnitudes range from 1.4 to 5 logits so that the categories are distinct and without large variance gaps between the categories.

Category fits are also calculated for each category, represented by the infit and outfit mean squares. If the outfit mean squares are greater than 2, it indicates that the category may introduce noise into the measurement process, which affects the quality of the rating scale.

2.2.3.6 Diagnosis (construct validity): Scale linearity

The linearity of a scale means that the degree of change in one part of the scale is equivalent to the same degree of change in another part of the scale. Scale linearity can be revealed by plotting the raw score on the y-axis against the item difficulty logits on the x-axis. In the graph, an ordinary least square regression line is plotted. The closer the curve is to the line, the more linear the scale (F. Wolfe et al., 2000).

2.2.3.7 Diagnosis (reliability): Item reliability

The item reliability index indicates the expected replicability of the item ordering if these items are administered to another sample of the same size, with the same traits. That is, with high item reliability, we would expect the item difficulty to remain stable if it is applied to another sample of stroke survivors with similar traits. There is a higher value of item reliability index if there is a wider range of item difficulty and a larger sample size. The high item reliability infers the consistency of the item difficulty hierarchy (Bond & Fox, 2007; Linacre, 2007).

2.2.3.8 Diagnosis (reliability): Person reliability

The person reliability index indicates the expected replicability of the ordering of the people if these people are administered another set of items measuring the same construct. That is, with a high person reliability index, we would expect a high probability that persons with high person ability logits will have higher measures on the second tool than persons with low person ability logits. A reliability value higher than .80 means that the more independent performers can be reliably distinguished from the less independent performers, while a reliability of .50 means that the performance differences may be due to random chance. Some factors affect the magnitude of the person reliability. The person separation reliability value is higher if there is (a) a wider range of ability in the sample, (b) a larger number of items in a test, (c) a longer rating scale, (d) a measurement tool with more categories per item, or (e) lower measurement error in the tool. The person reliability index is a ratio of the model-reproducible amount of the variance to the total variance of the person abilities, which is equivalent to the traditional test reliability that is the ratio of the true variance (the real variance under ideal conditions) to the total variance (including the true variance and the error variance) (Bond & Fox, 2007; Linacre, 2007; Portney & Watkins, 2000). The person reliability index is also analogous to the Cronbach's alpha (Linacre, 1997, 2007).

An alternative method to express reliability is the person separation index. This index indicates the stability of the person stratification levels of a measure. It is a ratio of the true standard deviation to the error standard deviation, or the root-mean-square standard error (RMSE). The value of the person separation index, analogous to a signal-to-noise ratio, is represented in RMSE units. Larger values of the person separation index imply a greater ability of the tool to differentiate person abilities. A person separation index higher than 2, corresponding to a person separation reliability of 0.80, implies good test reliability, and a 1, corresponding to a reliability of 0.50, implies the differences of the persons within the sample may be due to measurement error (Bond & Fox, 2007; Fisher Jr., 1992; Linacre, 2007; Wright, 1996; Wright & Masters, 1996).

The Cronbach's alpha is also calculated from the raw scores of the sample by WINSTEPS, along with the person separation reliability. It indicates the internal consistency of an instrument (Bond & Fox, 2007; Linacre, 2007; Portney & Watkins, 2000). If a value approaches 0.90, the instrument is considered internally consistent. Moderate consistency of an instrument is interpreted if a value is between 0.70 and 0.90 (Portney & Watkins, 2000).

2.2.3.9 Anchoring item values

An essential concept of measurement is that the calibration of the items in an instrument should remain invariant so that the measuring will remain valid, regardless of the intended purposes (Bond & Fox, 2007). To maintain the invariance of item calibration, before analyzing a subset of data, one can anchor the item values. For example, the anchored values of item difficulty logits, for the stroke survivor population, are estimated from a combined dataset which contains numerous sub-datasets collected in different contexts (i.e., data from patients' homes) and times (i.e., 3, 6, 9, and 12-months post-stroke in the major study). This step allows the item difficulty hierarchy to be fixed, and analyses of person abilities to be calculated for a specific population (e.g., stroke survivors at 3 months post-stroke). Thus, the values of person ability logits are calculated based on the anchored values of item difficulty logits (Bond & Fox, 2007; Linacre, 2007).

2.2.4 Correlations

After using Rasch analysis to transform PASS and BI ordinal data into interval data, person ability logits for 3 months post-stroke participants were established, using the anchored items. To test the hypotheses of the current study, the relationships of the scores between the instruments (i.e., logits of PASS and BI and original scores of mRS and GOS) were examined using Pearson Product-Moment correlations or Spearman's Rho correlations, as appropriate, using SPSS version 14.0. Significance was established at $p < .05$. The criteria used for evaluating the strength of the correlation coefficients were: (a) little or no relationship ranges from 0.00 to 0.25, (b) fair degree of relationship ranges from 0.25 to 0.50, (c) moderate to good relationship ranges from 0.50 to 0.75, and (d) a good to excellent relationship is greater than 0.75 (Portney & Watkins, 2000).

2.3 RESULTS

2.3.1 Participants

To anchor the item difficulty of a measure and to diagnose the construct of a measure using Rasch analysis, the dataset that contained data of the target population (i.e., stroke survivors) should be analyzed. In the current study, there were up to 213 administrations using the five ADL measures, including testing phases of 3, 6, 9, and 12-months post-stroke. The characteristics of the dataset are presented in Table 2-6. The sample was predominantly male (58.2%), white (97.7%), and had a mean age of 63.35 years. 52.7% of the sample was diagnosed with left hemispheric stroke (LHS), 38.2% with right hemispheric stroke (RHS), 6.3% with cerebellum stroke, and 2.9% with brainstem stroke. Participants with left hemispheric stroke constituted 57.4% of the sample. Other data from medical histories or factors related to stroke are also summarized in this table.

The characteristics of the 68 participants at 3 months post-stroke are presented in Table 2-7. The sample was predominantly male (63.2%), white (97.1%), and had a mean age of 65.53 years. The participants with ischemic stroke constituted 85.3% of the sample. Participants with left hemispheric stroke constituted 57.4% of the sample. There was no significant difference in the number of participants with left and right hemispheric strokes (LHS and RHS) ($\chi^2 = 1.47$, 1df, $p = 0.23$). Other data from medical histories or factors related to stroke are also summarized (see Table 2-6).

Table 2-6: Participants' Demographics and Stroke Characteristics for the Anchoring and 3 Month Post-Stroke Datasets

	Anchoring (N = 213)	Three Months (N = 68)
Gender		
Male (%)	58.2	63.2
Age, years (M, *SD*)	63.35 (15.24)	65.53 (14.03)
Ethnicity		
White (%)	97.7	97.1
Black (%)	1.9	1.5
Other (%)	0.5	1.5
Stroke location		
Left Hemisphere (%)	52.7[a]	57.4[c]
Right Hemisphere (%)	38.2[a]	42.6[c]
Cerebellum (%)	6.3[a]	N/A
Brainstem (%)	2.9[a]	N/A
Stroke type		
Ischemic (%)	88.3	85.3
Hemorrhagic (%)	11.7	14.7
Has prior CVA (%)	6.3[b]	7.4[d]
Has history of hypertension (%)	57.7[b]	63.2[d]
Has history of cardiac medication use (%)	35.1[b]	42.6[d]
Has history of diabetes mellitus (%)	23.1[b]	26.5[d]
Current cigarette smoker (%)	23.1[b]	23.5[d]
History of alcohol abuse (%)	6.3[b]	5.9[d]
History of atrial fibrillation (%)	13.5[b]	13.2[d]
Has used antiplates medication in week prior to stroke (%)	29.3[b]	35.3[d]

Note. [a]n = 207 for stroke location. [b]n = 208 for each characteristic. [c]No significant difference ($\chi^2_{(1)} = 1.47$, $p = .23$) between the percentages of the participants with left versus right hemispheric stroke. [d]n = 66 for each characteristic.

2.3.2 Category analysis of global non-summative measures

Table 2-7 summarizes the category responses of each category in the GOS5. In GOS5, the least observed response was a score of 4. The observed percentage did not increase as the value of the scores increased, or the functional levels increased. Scores 1 and 2 were absent for the sample, indicating that these two categories may not represent functional status of stroke participants at 3 months post stroke. The histogram for the category frequency of the GOS5 is shown in Figure 2-1. The distribution curve was skewed to the left. Table 2-8 summarizes the category responses of each category in the GOS8. The least observed responses were scores of 6 and 8. Again, the observed percentage did not increase as the value of the scores increased, or the functional levels increased. Scores 1 and 2 were, again, absent for the sample, indicating scores 1, 2, 6, and 8 may not represent functional status of stroke participants at 3 months post-stroke. Figure 2-2 represents the histogram for the category frequency of the GOS8. The distribution curve had a slight negative skew. Table 2-9 summarizes the category responses of each category in the mRS. The least observed responses were scores 0 and 5. The observed percentage did not increase as the value of the scores decreased, or the functional levels improved, but its distribution was normal (see Figure 2-3).

Table 2-7: Category Responses of Glasgow Outcome Scale, 5-point version (GOS5) (n = 68)

	Score	Frequency	%
-	1	0	0.0
	2	0	0.0
	3	34	50.0
	4	15	22.1
+	5	19	27.9

Note. % = Percentage of the responses for each category. + = The category represents higher function of a participant than other categories. - = The category represents lower function of a participant than other categories.

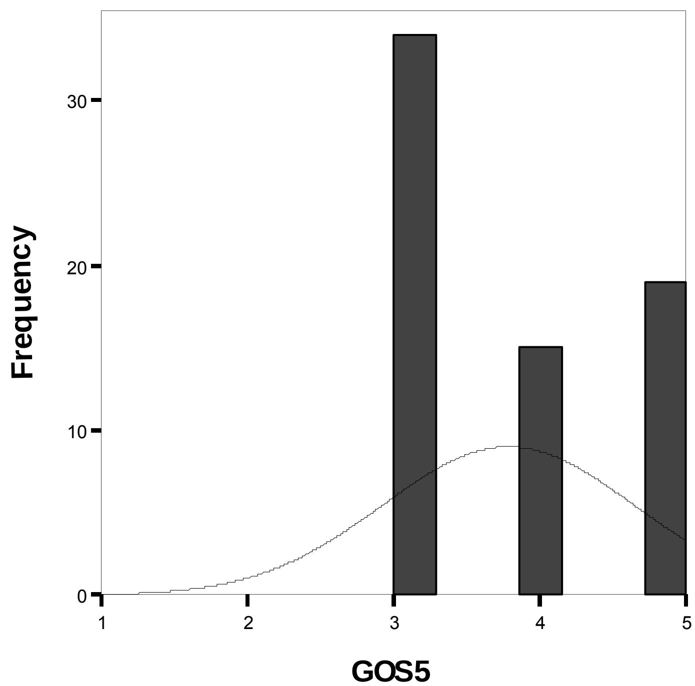

Figure 2-1: Histogram of the category frequency of the Glasgow Outcome Scale, 5-point version (GOS5).
The curved line represents the shape of the distribution (n = 68).

Table 2-8: Category Responses of Glasgow Outcome Scale, 8-point version (GOS8) (n = 68)

	Score	Frequency	%
-	1	0	0.0
	2	0	0.0
	3	12	17.6
	4	22	32.4
	5	14	20.6
	6	1	1.5
	7	18	26.5
+	8	1	1.5

Note. % = Percentage of the responses for each category. + = The category represents higher function of a participant than other categories. - = The category represents lower function of a participant than other categories.

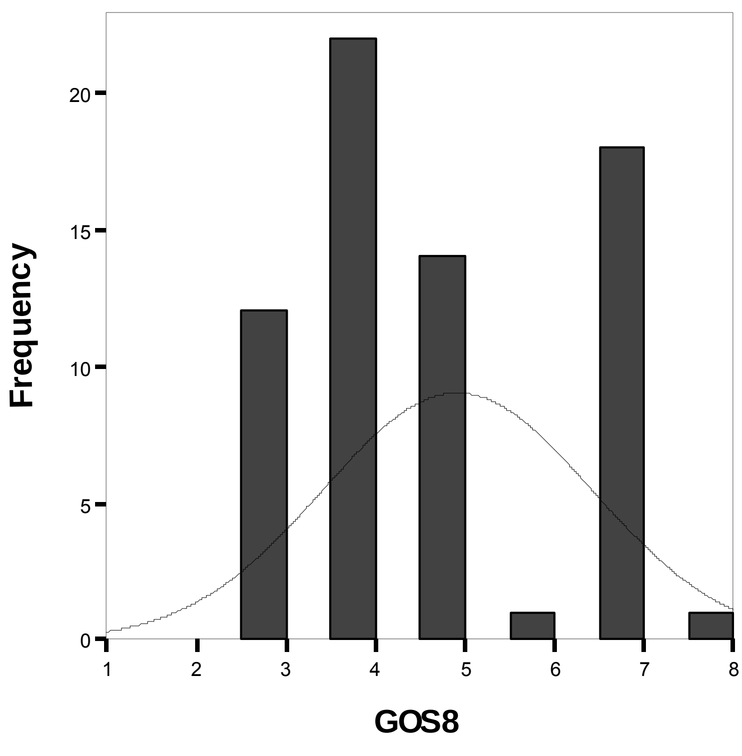

Figure 2-2: Histogram of the category frequency of the Glasgow Outcome Scale, 8-point version (GOS8).
The curved line represents the shape of the distribution (n = 68).

Table 2-9: Category Responses of Modified Rankin Scale (mRS) (n = 68)

	mRS scores	Frequency	%
+	0	3	4.4
	1	15	22.1
	2	12	17.6
	3	28	41.2
	4	7	10.3
-	5	3	4.4

Note. % = Percentage of the responses for each category. + = The category represents higher function of a participant than other categories. - = The category represents lower function of a participant than other categories.

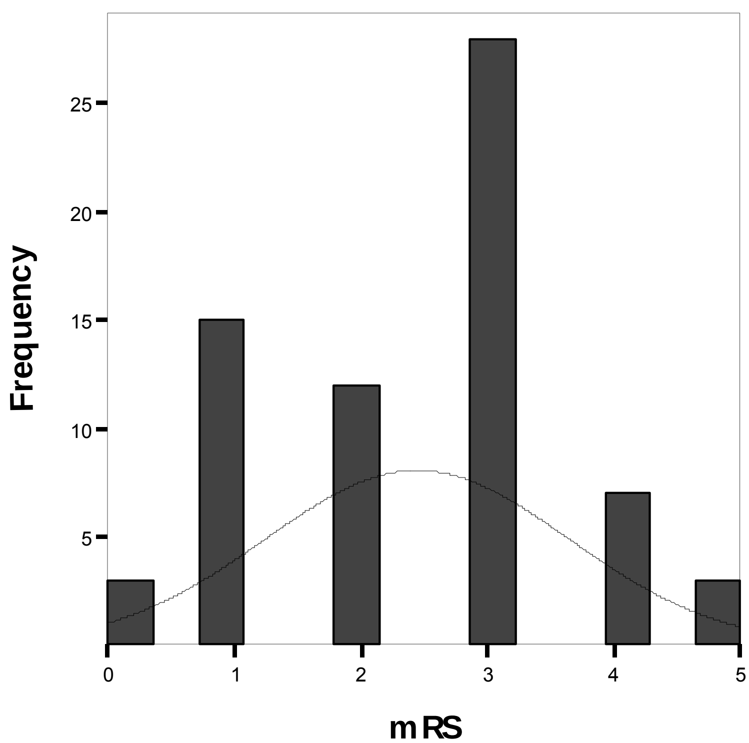

Figure 2-3: Histogram of the category frequency of the Modified Rankin Scale (mRS). The curved line represents the shape of the distribution (N = 68).

2.3.3 Rasch analysis

2.3.3.1 Diagnosis of construct validity: Barthel Index (BI)

BI-Fit statistics

Rasch analysis was used to analyze 213 administrations of the BI. Table 2-10 summarizes the anchored item difficulty logits and fit statistics of BI items for the stroke participants. Two items, toilet use and feeding, had values of both infit and outfit statistics ≤ 0.5 and ≥ 1.7. They were further inspected because the Rasch model did not predict these responses and the responses in reality may not be equally distributed along the continuum (i.e., independence or dependence). The misfit items also indicated that the data from the BI was not an excellent fit with the linear logit scale, and the unidimensionality may not be excellent, so these would be examined further in the principal component analysis. Among the 213 administrations for the toilet use item, 8.9% reported dependence, 8% required some assistance (e.g., imbalance or other problems with clothes or toilet paper), and 83.1% were independent in all actions (getting on and off toilet, fastening and unfastening clothes, using toilet paper without assistance, using wall bar or other support if needed, and using bedpan appropriately if necessary). For the feeding items, 6% reported dependence, 45 % required some assistance (e.g., cutting food or spreading butter), and 49% were independent in all actions (feeding self from tray or table, putting on assistive device if needed, and accomplishing feeding in reasonable time). Although the items were not normally distributed, they were consistent with performance of the population, and thus they were kept in the model.

BI – Principal component analysis of Rasch residuals

The standard residual variance scree plot of the principal component analysis (PCA) of Rasch residuals for the Barthel Index (BI) is presented in Figure 2-4. The Rasch model explained 82.0% of the variance out of the total variance (100%). If the data fit the Rasch model perfectly, the measurement dimension would explain 84.3% of the variance. The unexplained variance for the BI was 18.0%, and the first contrast in the residuals explained 3.5% of the variance. The eigenvalue of the first contrast was 1.9, indicating that it had the strength of about 2 items out of 10 items in this analysis. The Rasch factor analysis for the first contrast yielded factor loadings, ranging from -0.58 to 0.66 (see Table 2-11). The opposed poles of the factor loadings distinguished the common variance of the items. The "A" (feeding) and "a" (toilet use) in the plot identify the items with the most opposed loadings on the first contrast in the residuals of the BI. The items with positive values of factor loadings included feeding, bladder, bowls, and grooming, which all involve upper body movements and autonomic nervous system functions. In contrast, the items with negative values of factor loadings include toilet use, stairs, bathing, dressing, transferring, and walking, which involve whole body movements. The factor loading against the item difficulty logits of each task in the BI were also plotted (see Figure 2-5), in which the plot represented a random pattern. The items with positive factor loadings generally were easier for the stroke survivors than those with negative factor loadings. Because there was no specific group of clustered items at a certain difficulty level in the plot, further investigations for the variance distributions were not required. In addition, the item difficulty logits yielded from the primary Rasch analysis explained 45.5/55.5 units of the variance and the first contrast accounted for 1.9 units. The factor sensitivity ratio was 0.0417, indicating that after the Rasch measure was extracted, the first contrast influenced 4% of the stability of the unexplained variance in the BI. That is, 4% of the unexplained variance in the BI was associated with the second largest dimension. Clinically, this impact was not large enough to influence how to measure or interpret the ADL functional status of stroke survivors. In summary, the construct of the BI that measured ADL independence of the stroke survivors was considered unidimensional.

Table 2-10: Fit Statistics and Rasch Item Difficulty Logits of the Barthel Index (BI) Items

BI items	Measure	Infit MNSQ	Outfit MNSQ	Raw
Bowels	-1.17	1.24	3.17	155
Bladder	-0.79	1.29	2.29	144
Grooming	0.58	1.00	0.75	60
Toilet use	0.03	0.43	0.29	133
Feeding	-0.40	1.97	1.99	126
Transfer (bed-chair)	-0.87	0.74	0.35	226
Walking	-0.79	0.72	0.68	217
Dressing	0.44	0.76	0.77	116
Stairs	1.17	0.70	0.69	107
Bathing	1.80	0.67	0.57	44
Mean (SD)	0.00 (.93)			

Note. Measure = Item difficulty logits. MNSQ = Mean Square. Raw = Sum of the scored responses to an item.

```
       STANDARDIZED RESIDUAL VARIANCE SCREE PLOT
Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)
                                                   Empirical       Modeled
Total variance in observations     =         55.5 100.0%          100.0%
Variance explained by measures     =         45.5  82.0%           84.3%
Unexplained variance (total)       =         10.0  18.0% 100.0%    15.7%
Unexplned variance in 1st contrast =          1.9   3.5%  19.2%
Unexplned variance in 2nd contrast =          1.6   2.8%  15.6%
Unexplned variance in 3rd contrast =          1.4   2.5%  13.8%
Unexplned variance in 4th contrast =          1.1   2.0%  11.1%
Unexplned variance in 5th contrast =          1.0   1.9%  10.4%

      VARIANCE COMPONENT SCREE PLOT
       +--+--+--+--+--+--+--+--+--+
   100%+  T                        +
       |     M                     |
  V 63%+                           +
  A    |                           |
  R 40%+                           +
  I    |                           |
  A 25%+                           +
  N    |                           |
  C 16%+        U                  +
  E    |                           |
    10%+                           +
  L    |                           |
  O  6%+                           +
  G    |                           |
  |  4%+                           +
  S    |           1               |
  C  3%+              2            +
  A    |                 3  4      |
  L  2%+                       5   +
  E    |                           |
  D  1%+                           +
       |                           |
   0.5%+                           +
       +--+--+--+--+--+--+--+--+--+
          TV MV UV U1 U2 U3 U4 U5
         VARIANCE COMPONENTS
```

Figure 2-4: Standard residual variance scree plot of principal component analysis (PCA) for the Barthel Index (BI).

T, TV represents the total variance in the observations. M, MV represents variance in the observations explained by the Rasch measures. U, UV represents unexplained variance. 1, U1 represent first contrast (component) in the residuals. 2 (U2), 3 (U3), 4 (U4) and 5 (U5) represent second, third, fourth and fifth contrasts (components) in the residuals, accordingly.

Table 2-11: Matrix of Standardized Residual Contrast 1 of PCA for BI.

Entry	BI items	Loading	Measure
A	Feeding	0.66	-0.40
B	Bladder	0.37	-0.79
C	Bowels	0.34	-1.17
D	Grooming	0.33	0.58
a	Toilet use	-0.58	0.03
b	Stairs	-0.56	1.17
c	Bathing	-0.51	1.80
d	Dressing	-0.33	0.44
e	Transfer (bed to chair)	-0.27	-0.87
E	Walking	-0.14	-0.79

Note. Measure = Item difficulty logits.

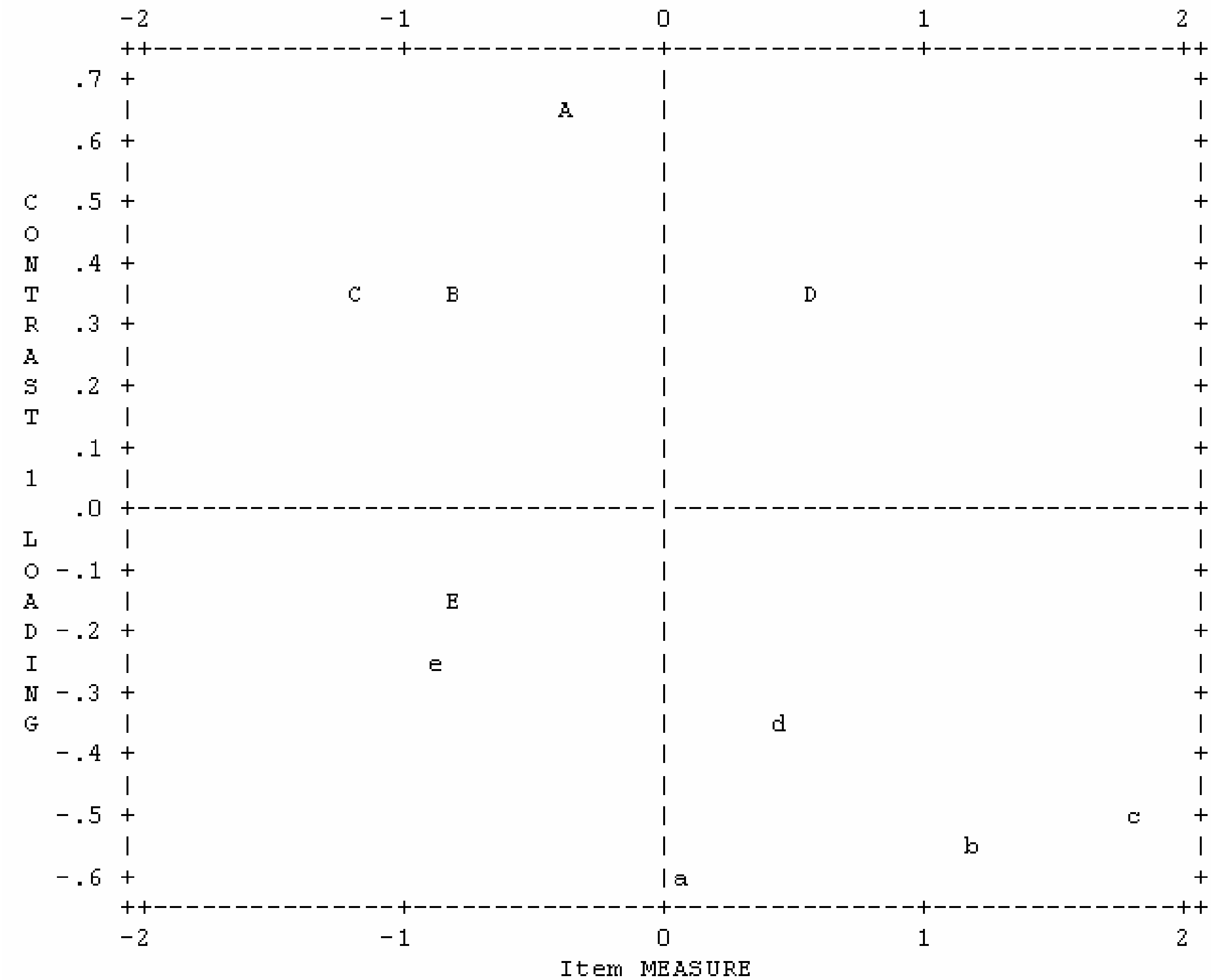

Figure 2-5: Standardized residual contrast 1 plot of principal component analysis (PCA) for the Barthel Index (BI).

The plot is plotted by the contrast 1 loadings and item measures (item difficulty logits) in the matrix of standardized residual contrast 1. This plot presents the clustered residuals visually. The “A” (feeding) and “a” (toilet use) in the plot identify the item loadings that are the most opposed on the first contrast in the residuals.

BI – Category function

Because the partial credit model was applied to the BI data, Rasch analysis category function of each item is presented (see Table 2-12). The observed percentage represents the rating frequency for each category. According to these numbers, the categories for five items (bowels, bladder, feeding, transfer, walking) may be problematic, because some categories had less than 10 responses for these items. The average person ability measures are shown by the "OBSVD AVERGE." For each item, the average measures increased in size as the variable increased, indicating that, on average, the persons with better abilities favored the higher category. The structure calibrations represent the thresholds of categories for each item, except two items with dichotomous choices (grooming and bathing). The structure calibrations of each item presented logical patterns such that categories with higher category levels were more difficult to obtain, except items for the bowels and walking. The score 1 in the bowels item had abnormally few observations, but was calibrated higher than a score of 2. Finally, the category fits were shown by the Infit and Outfit mean squares. Most category fit statistics were within the acceptable range, less than 2. Specifically, the categories for item feeding showed misfit statistics, indicating that they introduced noise into the measurement process. Overall, the categories of the BI functioned moderately.

Table 2-12: Category Functions for Items of the Barthel Index (BI).

Item	Category	Observed %	OBSVD AVRGE	Structure calibration	Infit MNSQ	Outfit MNSQ
Bowels						
	0	9	-2.11	None	0.99	0.48
	1	6	1.46	0.49	1.53	4.52
	2	85	1.99	-0.49	1.41	1.15
Bladder						
	0	8	-2.40	None	0.50	0.41
	1	20	1.50	-0.99	1.47	2.94
	2	72	2.06	0.99	1.71	1.52
Grooming						
	0	32	-0.10	N/A	0.97	0.65
	1	68	2.38	N/A	1.05	0.98
Toilet use						
	0	15	-1.76	None	0.39	0.30
	1	19	0.62	-0.49	0.47	0.18
	2	66	2.62	0.49	0.45	0.58
Feeding						
	0	6	- 0.21	None	2.38	2.01
	1	45	1.44	-2.25	1.68	1.84
	2	49	1.94	2.25	2.16	2.19
Transfer						
	0	8	-2.44	None	0.58	0.44
	1	5	-0.98	-0.17	0.51	0.13
	2	10	0.31	-0.15	0.63	0.25
	3	77	2.32	0.32	1.13	1.12
Walking						
	0	7	-2.66	None	0.56	0.43
	1	2	-0.98	0.13	0.70	0.44
	2	28	0.69	-1.86	0.81	0.67
	3	63	2.56	1.73	0.71	0.87
Dressing						
	0	13	-1.86	None	0.51	0.52
	1	43	1.35	-1.64	0.86	0.83
	2	44	2.79	1.64	0.83	0.84
Stairs						
	0	27	-0.48	None	0.77	0.61
	1	24	1.23	-0.35	0.88	0.85
	2	49	2.92	0.35	0.52	0.57
Bathing						
	0	50	0.32	N/A	0.68	0.59
	1	50	2.86	N/A	0.66	0.54

Note. Observed % = Percentage of occurrences of that category. OBSVD AVERGE = Average of the person ability logits for all persons in the sample who had that particular response to that category. Structure calibration = The difficulties estimated for choosing one response category over another. MNSQ = Mean square.

BI – Scale linearity

The scale linearity of the BI is shown in Figure 2-6. The raw data were the sum of the scored responses to each item (see Table 2-10) and the measures were the item difficulty logits of each item. The R^2 linearity value is 0.663, about 66% of the variance can be explained by the model (Portney & Watkins, 2000). It indicates that the scale of the BI is somewhat linear.

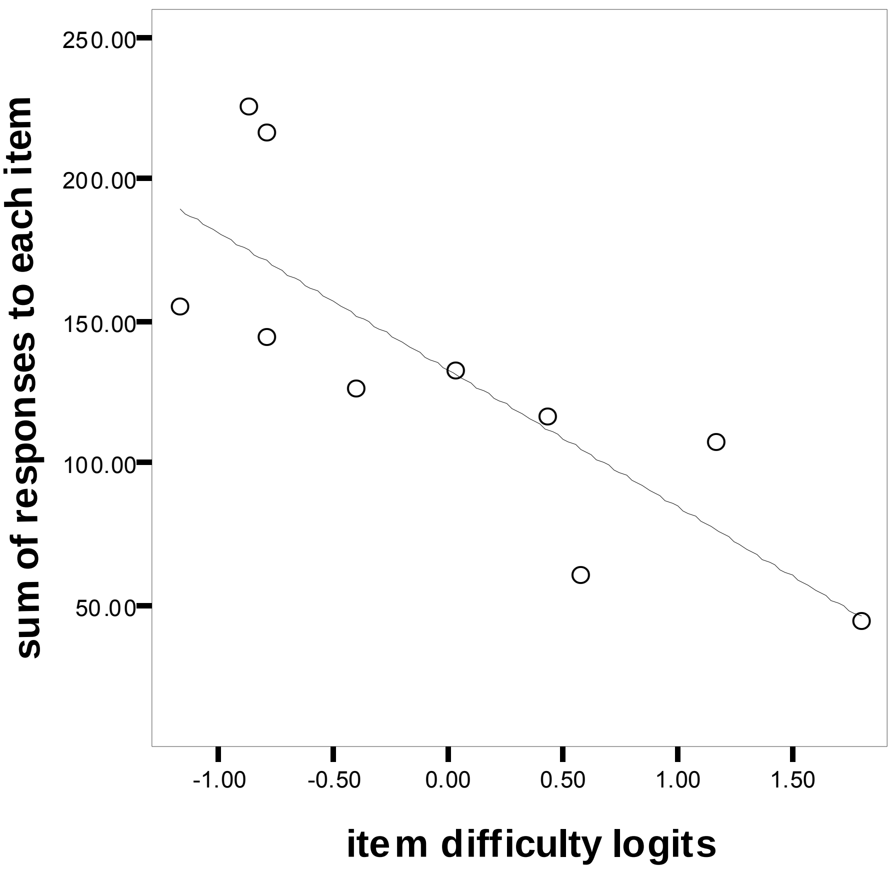

Figure 2-6: Scale linearity of the Barthel Index (BI).
Each dot represents one item of the BI. They were plotted by their item difficulty logits (x-axis) and sum of responses to each item in the Rasch analysis in the Rasch analysis (y-axis). The line was ordinary least squares regression of y on x.

2.3.3.2 Diagnosis of reliability: Barthel Index (BI)

BI – Person reliability

The person separation reliability index of the BI was 0.78, approaching good reproducibility. It also implied that the persons who were more independent could be moderately distinguished from those who were less independent by the BI. The person separation index of the BI was 1.87, which is also approaching good reproducibility of the person ordering. In addition, the internal consistency of the BI was excellent with a Cronbach's Alpha of 0.95. The above data showed that the BI is a moderately reliable tool to evaluate ADL performance for stroke survivors.

BI – Item reliability

The item reliability index of the BI was 0.92, which indicated that the item difficulty hierarchy of the BI would be consistent over other populations with the same traits and same sample size as our data.

2.3.3.3 Person ability logits of the Barthel Index (BI)

Figure 2-7 presents the item-person map for the BI. The map illustrates the distributions of the abilities of the persons at 3 months stroke, along the ability logit scale, and the difficulties of the BI items, along the difficulty logit scale, at the same time. The distribution of the persons was not as clumped as for the BI items. Figure 2-8 was further plotted to show the person ability logits distribution of the BI. The curve in the graph was skewed to the left, indicating that a majority of the sample showed high ability in the BI.

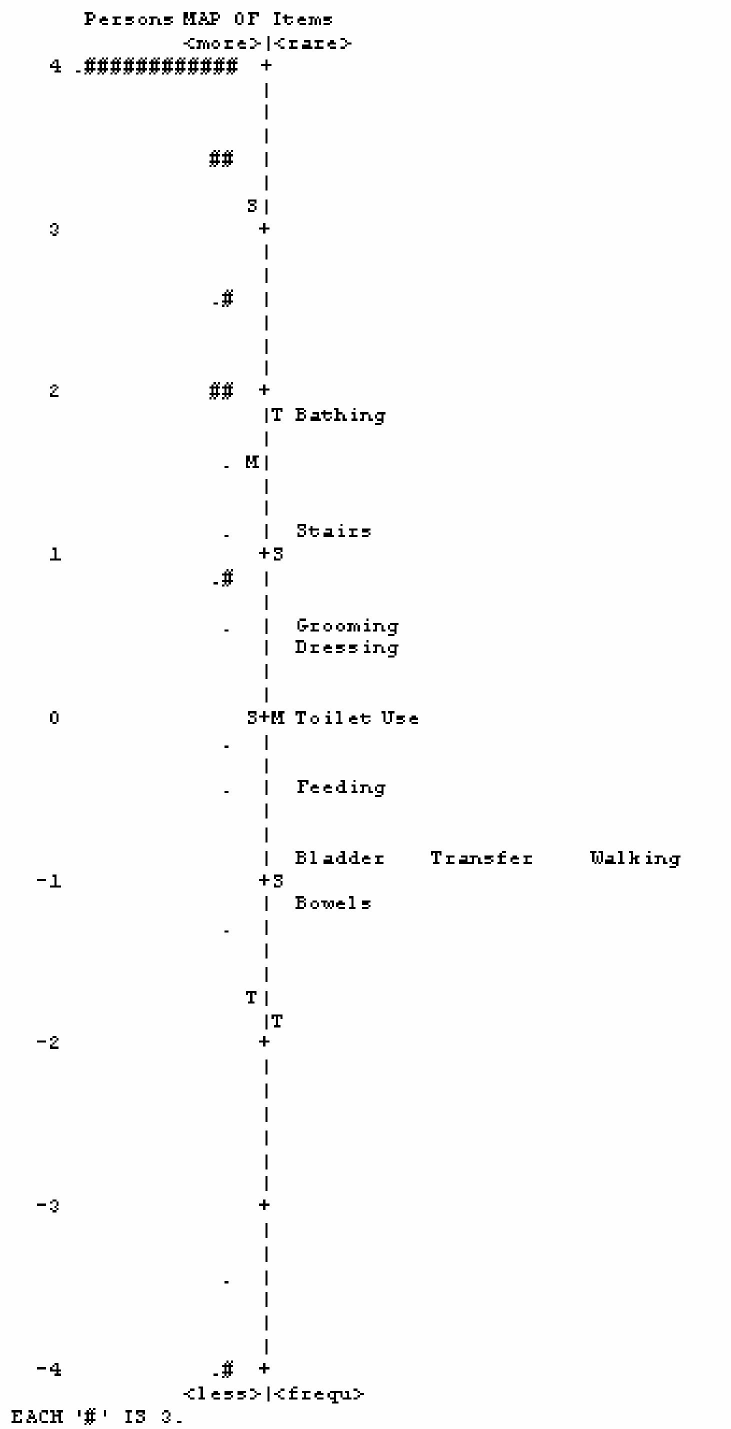

Figure 2-7: Item-person map for the Barthel Index (BI).
To the left of the dashed line (the logit scale) is the person ability scale, distributed from the greatest ability at the top to the least ability at the bottom. Each # represents three people in the sample. To the right of the dashed line is the item difficulty scale distributed from the most difficult item at the top to the easiest item at the bottom. Along the logit scale, "M" represents the mean person ability or item difficulty estimate, "S" represents the location of one standard deviation (SD) from the mean estimates, and "T" is the second SD away from the mean estimates.

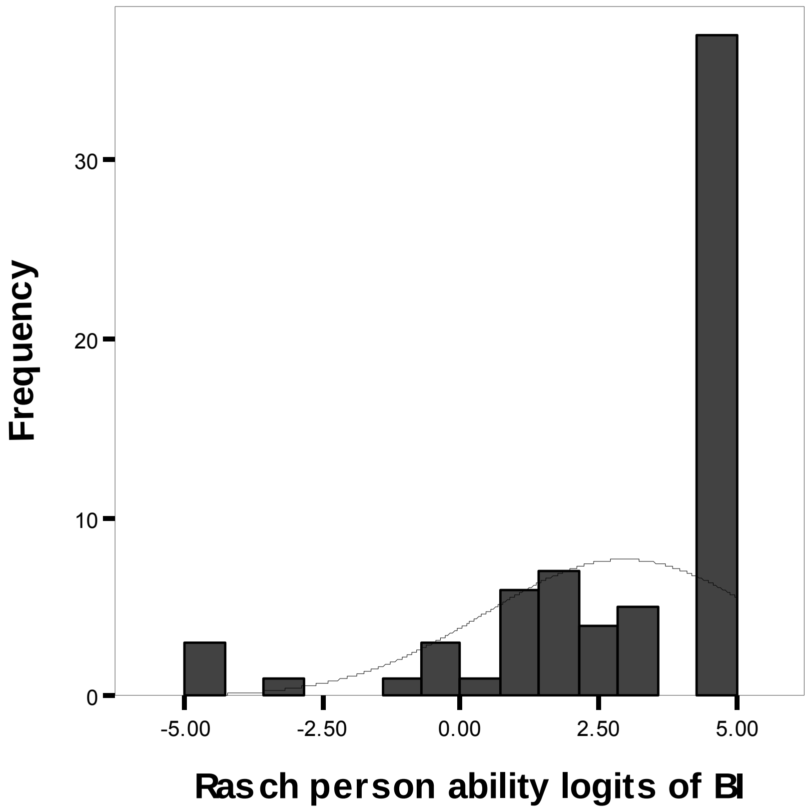

Figure 2-8: Histogram of the person ability logits frequency for the Barthel Index (BI). The curved line represents the shape of the distribution (N = 68).

2.3.3.4 Diagnosis of construct validity: Performance Assessment of Self-Care Skills (PASS)

PASS – Fit statistics

Rasch analysis was used to analyze 211 administrations of the PASS. Table 2-13 summarizes the anchored item difficulty logits and fit statistics of PASS items for the stroke participants. While some values of the outfit mean squares were out of the recommended range, no task had both fit statistics values ≤ 0.5 and ≥ 1.7.

PASS – Principal component analysis of Rasch residuals

The standard residual variance scree plot of principal component analysis (PCA) for the PASS is presented in Figure 2-9. The Rasch model explained 91.7% of the total variance (100%). If the data fit the Rasch model perfectly, the measurement dimension would explain 91.5% of the variance. The unexplained variance for PASS was 8.3%, and the first contrast in the residuals explained 1.5% of the variance. The eigenvalue of the first contrast was 4.5, indicating that it had the strength of about 4 or 5 items of 25 items in this analysis. Based on the rules of thumb of PCA, the construct of assistance in the PASS, for the stroke participants, was unidimensional. The Rasch factor analysis for the first contrast yielded factor loadings, ranging from -0.62 to 0.72 (see Table 2-14). The opposed poles of the factor loadings distinguished the common variance of the items. The "A" (checkbook balancing) and "a" (toilet transfers) in the plot identified the items with the most opposed loadings on the first contrast in the residuals of the PASS. The tasks with positive values of factor loadings included checkbook balancing, bill paying, mailing, obtaining visual information, shopping, medication management, obtaining auditory information, playing bingo, home safety, and telephone use, all of which are cognitive and table-top activities. In contrast, the tasks with negative values of factor loadings included toilet transfers, indoor walking, bed transfers, sweeping, dressing, oral hygiene, stair use, carrying garbage, cleanup after meal preparation, changing bed linens, stovetop use, bathtub or shower transfers, using sharp utensils, small repairs, and trimming toenails, all of which require greater physical motion and interaction with the environment than the tasks with positive factor loadings. The factor loading against the item difficulty logits of each task in the PASS was also plotted (see Figure 2-10), in which the plot represented a random pattern. The items with negative factor loadings generally had a wider distribution of item difficulty levels than those with positive factor loadings. Because there was no specific group of clustered items at a certain difficulty level in the plot, further investigations for the variance distributions were not required. In addition, the item difficulty logits yielded from the primary Rasch analysis explained 276.1/301.1 units of the variance and the first contrast accounted for 4.5 units. The factor sensitivity ratio was 0.0162, indicating that after the Rasch measure was extracted, the first contrast influenced about 2% of the unexplained variance in the PASS. That is, 2% of the unexplained variance in the PASS was associated with the second large dimension. Clinically, this impact was not large enough to influence how to measure or interpret the ADL functional status of stroke survivors. In summary, the construct of the PASS that measured ADL independence of the stroke survivors was considered unidimensional.

Table 2-13: Fit Statistics and Rasch Item Difficulty Logits of the Performance Assessment of Self-care Skills (PASS) Task Items

PASS tasks	Logits	Infit MNSQ	Outfit MNSQ	Raw
Bed transfers	-1.36	0.85	0.95	511
Stair use	-0.33	1.24	1.36	432
Toilet transfers	-1.20	0.83	0.64	503
Oral hygiene	-1.22	0.66	0.47	504
Bathtub/Shower transfers	-0.29	1.15	1.77	427
Trimming toenails	1.27	1.48	1.29	194
Dressing	-0.62	0.77	0.73	461
Shopping	0.34	0.73	0.84	341
Bill paying	0.68	0.89	0.90	286
Checkbook balancing	0.92	1.00	1.04	248
Mailing	0.96	0.77	0.82	242
Carrying garbage	0.67	1.41	1.38	288
Telephone use	0.01	0.91	0.59	388
Medication management	0.02	0.87	1.43	387
Changing bed linens	1.04	1.31	1.35	229
Obtaining information-auditory	-0.52	1.28	1.36	452
Obtaining information-visual	-0.21	1.39	1.53	417
Small repairs	-0.28	0.87	0.70	426
Sweeping	0.27	1.10	0.82	351
Indoor walking	-1.17	1.27	0.95	501
Home safety	-0.41	0.87	0.66	440
Playing bingo	-0.25	1.21	1.56	422
Stovetop use	0.55	0.63	0.43	307
Using sharp utensils	0.64	0.83	0.60	293
Cleanup after meal preparation	0.49	0.65	0.43	317
Mean (SD)	0.00 (.74)			

Note. Measure = Item difficulty logits. MNSQ = Mean Square. Raw = Sum of the scored responses to an item.

```
       STANDARDIZED RESIDUAL VARIANCE SCREE PLOT
Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)
                                                 Empirical       Modeled
Total variance in observations     =       301.1 100.0%          100.0%
Variance explained by measures     =       276.1  91.7%           91.5%
Unexplained variance (total)       =        25.0   8.3% 100.0%     8.5%
Unexplned variance in 1st contrast =         4.5   1.5%  17.9%
Unexplned variance in 2nd contrast =         2.6    .9%  10.4%
Unexplned variance in 3rd contrast =         2.1    .7%   8.5%
Unexplned variance in 4th contrast =         1.9    .6%   7.8%
Unexplned variance in 5th contrast =         1.4    .5%   5.7%

      VARIANCE COMPONENT SCREE PLOT
       +--+--+--+--+--+--+--+--+--+
   100%+  T                       +
       |     M                    |
  V 63%+                          +
  A    |                          |
  R 40%+                          +
  I    |                          |
  A 25%+                          +
  N    |                          |
  C 16%+                          +
  E    |                          |
    10%+                          +
  L    |        U                 |
  O  6%+                          +
  G    |                          |
  |  4%+                          +
  S    |                          |
  C  3%+                          +
  A    |                          |
  L  2%+                          +
  E    |           1              |
  D  1%+                          +
       |              2           |
   0.5%+                 3  4  5  +
       +--+--+--+--+--+--+--+--+--+
          TV MV UV U1 U2 U3 U4 U5
         VARIANCE COMPONENTS
```

Figure 2-9: Standard residual variance scree plot of principal component analysis (PCA) for Performance Assessment of Self-care Skills (PASS).

T, TV represents the total variance in the observations. M, MV represents variance in the observations explained by the Rasch measures. U, UV represents unexplained variance. 1, U1 represent first contrast (component) in the residuals. 2 (U2), 3 (U3), 4 (U4) and 5 (U5) represent second, third, fourth and fifth contrasts (components) in the residuals, accordingly.

Table 2-14: Matrix of Standardized Residual Contrast 1 of PCA for the PASS.

	PASS tasks	Loading	Measure
A	Checkbook balancing	0.72	0.92
B	Bill paying	0.65	0.68
C	Mailing	0.65	0.96
D	Obtaining information-visual	0.53	-0.21
E	Shopping	0.46	0.34
F	Medication management	0.37	0.02
G	Obtaining information-auditory	0.35	-0.52
H	Playing bingo	0.31	-0.25
I	Home safety	0.21	-0.41
J	Telephone use	0.11	0.01
a	Toilet transfers	-0.62	-1.20
b	Indoor walking	-0.59	-1.17
c	Bed transfers	-0.51	-1.36
d	Sweeping	-0.43	0.27
e	Dressing	-0.42	-0.62
f	Oral hygiene	-0.39	-1.22
g	Stair use	-0.38	-0.33
h	Carrying garbage	-0.38	0.67
i	Cleanup after meal preparation	-0.35	0.49
j	Changing bed linens	-0.31	1.04
k	Stovetop use	-0.30	0.55
l	Bathtub/Shower transfers	-0.21	-0.29
M	Using sharp utensils	-0.18	0.64
L	Small repairs	-0.15	-0.28
K	Trimming toenails	-0.14	1.27

Note. Measure = Item difficulty logits.

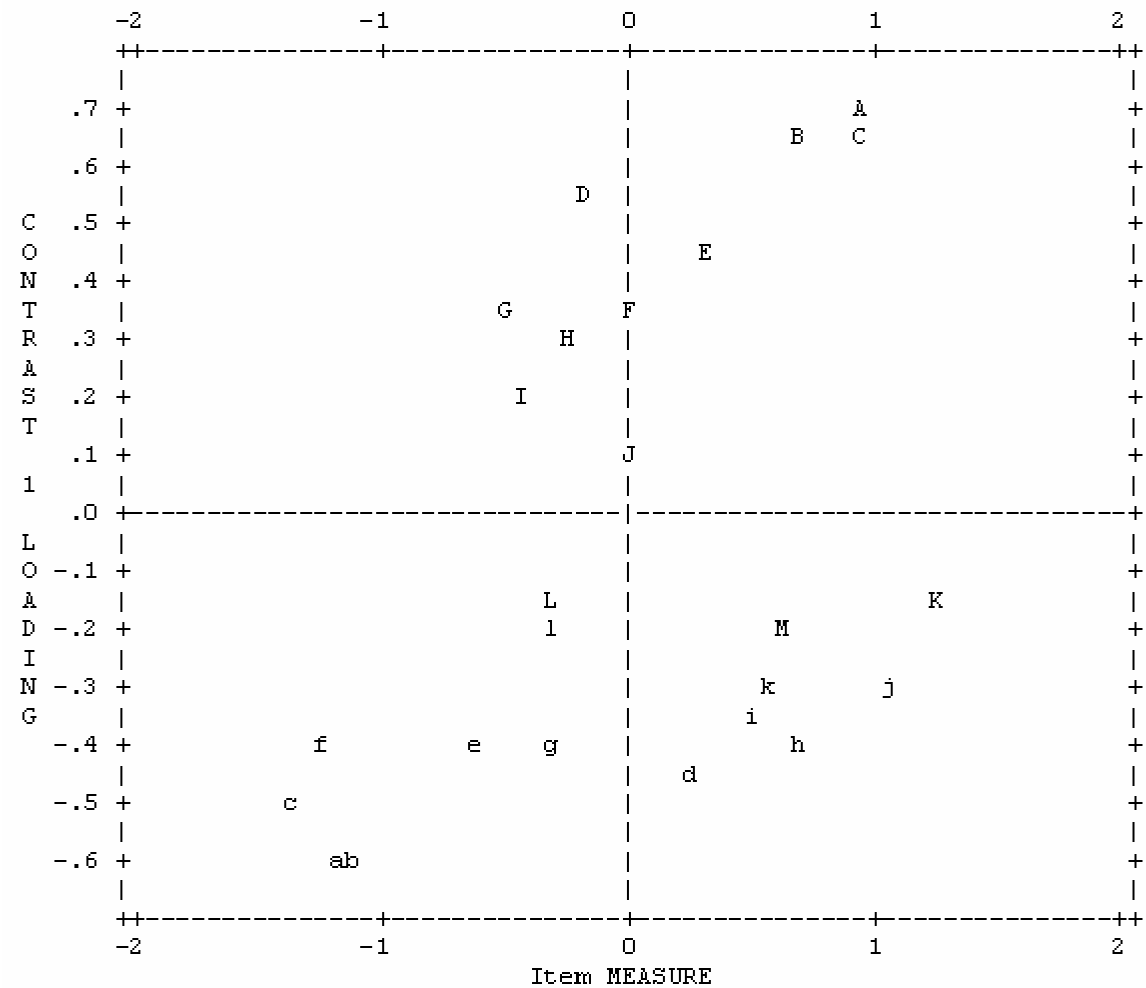

Figure 2-10: Standardized residual contrast 1 plot of principal component analysis for Performance Assessment of Self-care Skills (PASS).
The plot is plotted by the contrast 1 loadings and item measures (item difficulty logits) in the matrix of standardized residual contrast 1. This plot presents the clustered residuals visually. The "A" (balancing a checkbook) and "a" (toilet transfer) in the plot identify the item loadings that are the most opposed on the first contrast in the residuals.

PASS – Category function

Table 2-15 presents the category function of the PASS. The observed percentage represents the category frequency of each category. According to the data, category 1 (i.e., occasional physical assists to continuous verbal assists) had the least observations during the administrations. The average person ability measures are shown by the "OBSVD AVERGE." The average measures increased in size as the variable increased, indicating that, on average, the persons with better abilities favored the higher category. The structure calibrations represent the thresholds of categories. Category 1 was the most difficult category to be observed. Although the values of the thresholds did not increase with the category values, each threshold distance was within the recommended range of 1.4 to 5 logits. Figure 2-11 presents the probability curves for the categories in the PASS, which are plotted based on the values of the thresholds. In the graph, each category had its distinct peak. The curve of category 1 seemed flat, but it was still useful. This curve spanned a large portion of the variable and was not over-shadowed. Finally, the category fits were shown by the Infit and Outfit mean squares. All fit statistics were within the acceptable range, less than 2, indicating that there was no noise introduced into the measurement process. Overall, the categories of the PASS functioned well.

Table 2-15: Category Functions for Items of the Performance Assessment of Self-care Skills (PASS).

Category	Observed %	OBSVD AVRGE	Structure calibration	Infit MNSQ	Outfit MNSQ
0	26	-0.82	None	0.96	1.05
1	4	0.00	1.47	1.15	1.14
2	13	0.62	-0.95	0.91	0.79
3	57	1.41	-0.52	1.06	1.07

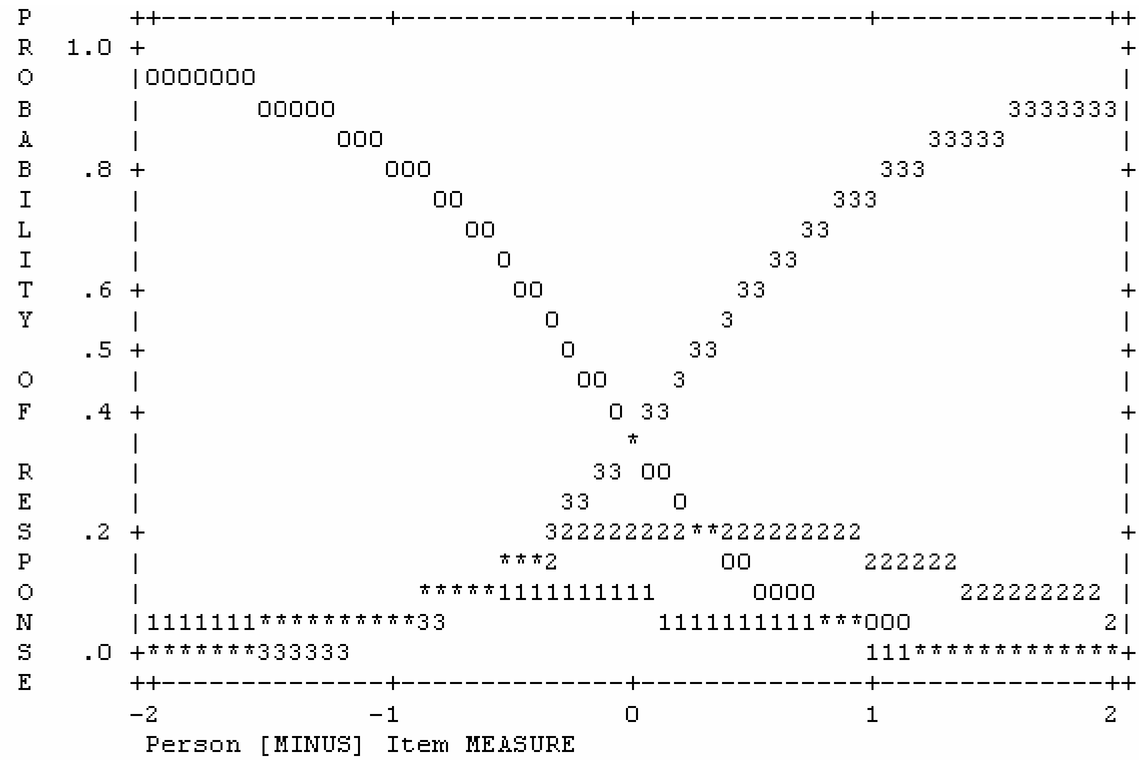

Figure 2-11: Probability curves for the Performance Assessment of Self-care Skills (PASS).

PASS – Scale linearity

The scale linearity of the PASS is shown in Figure 2-12. The raw data were the sum of the scored responses to each item (see Table 2-13) and the measures were the item difficulty logits of each item. The R^2 linearity value is 0.973, indicating that about 97% of the variance can be explained by the model (Portney & Watkins, 2000). It also indicates that the scale of the PASS is linear.

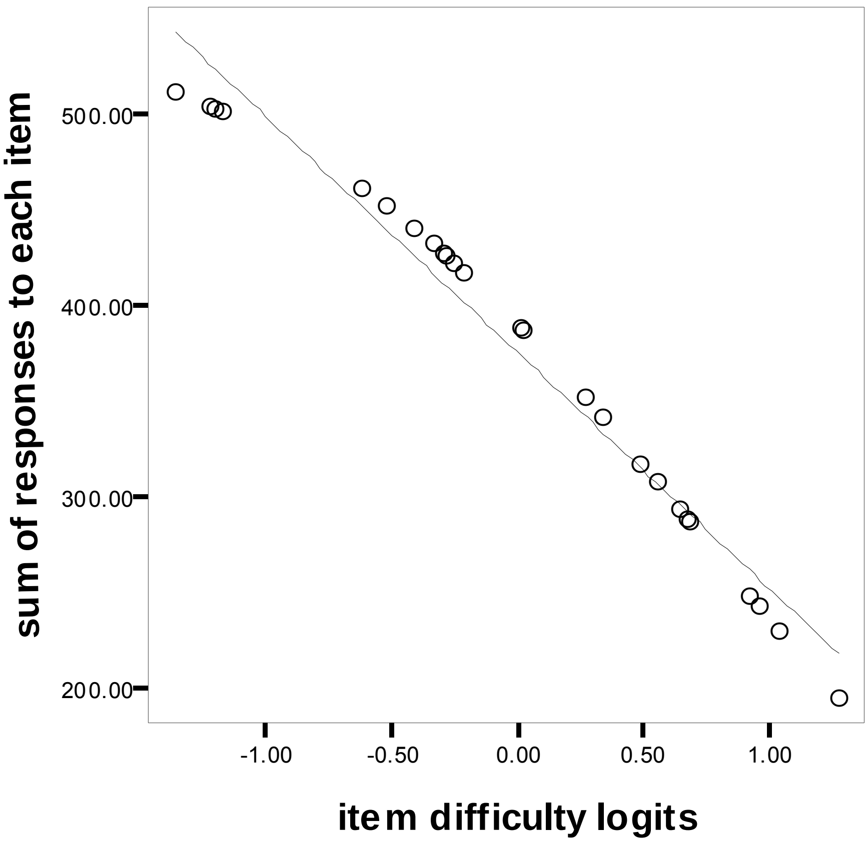

Figure 2-12: Scale linearity of the Performance Assessment of Self-care Skills (PASS).
Each dot represents one item of the PASS. They were plotted by their item difficulty logits (x-axis) and sum of responses to each item in the Rasch analysis in the Rasch analysis (y-axis). The line was the ordinary least squares regression of y on x.

2.3.3.5 Diagnosis of reliability: Performance Assessment of Self-Care Skills (PASS)

PASS – Item reliability

The item reliability index of the PASS was 0.98, which indicated the item difficulty hierarchy of the PASS would be consistent over other populations with the same traits and same sample size as our data. It also implied that the PASS is an excellent and valid tool to evaluate ADL performance for stroke survivors.

PASS – Person reliability

The person reliability index of the PASS was 0.89, indicating a good to excellent reproducibility. It also implied that the persons who were more independent could be reliably distinguished from those who were less independent by the PASS. The person separation index of the PASS was 2.84, which indicated an excellent stability of the person ordering. The internal consistency of the PASS was also excellent with a Cronbach's Alpha of 0.96. The above data showed that the PASS is an extremely reliable tool to evaluate ADL performance for stroke survivors.

2.3.3.6 Person ability logits of the Performance Assessment of Self-Care Skills (PASS)

Figure 2-13 presents the item-person map for the PASS. The map illustrates the distributions of the abilities of the persons at 3 months after stroke, along the ability logit scale, and the difficulties of the PASS task items, along the difficulty logit scale, at the same time. The distribution of the persons was clustered around the PASS task items, indicating that our sample found PASS tasks challenging. Figure 2-14 was further plotted to show the person ability logits distribution of the PASS. The curve in the graph was normally distributed, implying that the PASS is responsive to different levels of functional status of the stroke survivors at 3 months post-stroke.

```
     Persons MAP OF Items
       <more>|<rare>
 4        XX  +
              |
              |
              |
              |
              |
              |
         XXX  |
              |
              |
 3            +
             T|
              |
              |
           X  |
              |
              |
       XXXXX  |
              |
              |
 2         X  +
              |
             S|
              |
          XX  |
          XX  |T
          XX  |
              |  teonail
           X  |
           X  |
 1        XX  +  linens    mail
        XXXX  |  check
          XX  |
       XXXXX  |S bend      bill
         XXX M|  stovetop  utensil
          XX  |  clean
           X  |
           X  |  sweep     shop
           X  |
          XX  |
 0      XXXX  +M tel       med
           X  |
          XX  |  vision    bingo
      XXXXXX  |  stair     tub        repair
           X  |  safety
           X S|  aud
              |  dress
         XXX  |S
              |
           X  |
-1            +
              |
              |  walk      toilet     oral
           X  |
           X  |  bed
              |T
              |
           X T|
           X  |
              |
-2     XXXXX  +
       <less>|<frequ>
```

Figure 2-13: Item-person map for the Performance Assessment of Self-care Skills (PASS).
To the left of the dashed line (the logit scale) is the person ability scale, distributed from the greatest ability at the top to the least ability at the bottom. Each X represents a person in the sample. To the right of the dashed line is the item difficulty scale distributed from the most difficult item at the top to the easiest item at the bottom. Along the logit scale, “M” represents the mean person ability or item difficulty estimate, “S” represents the location of one standard deviation (SD) from the mean estimates, and “T” is the second SD away from the mean estimates.

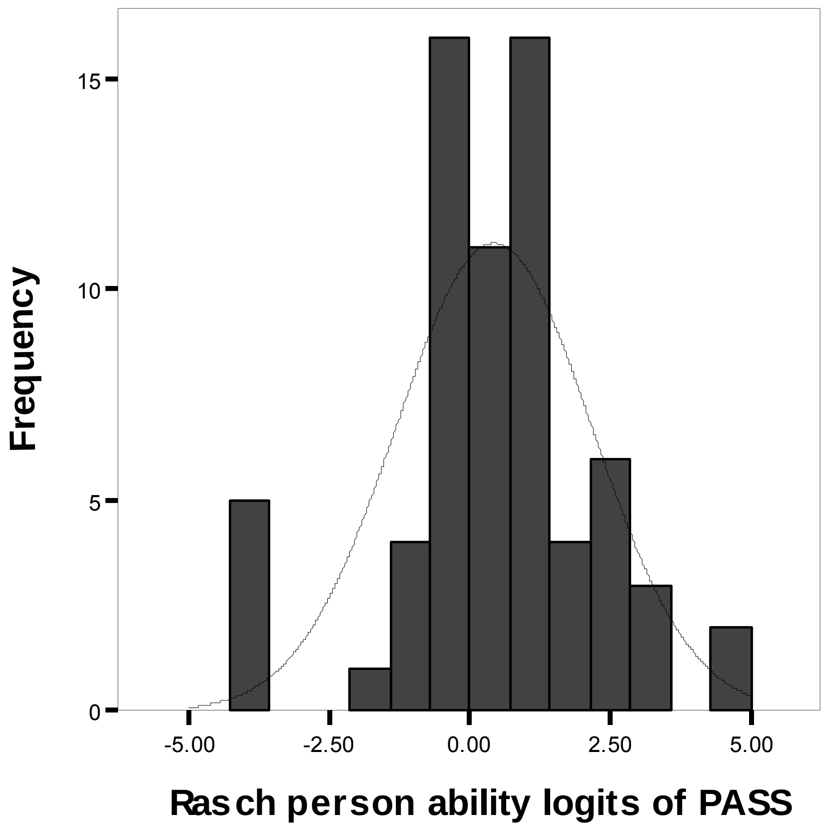

Figure 2-14: Histogram of the person ability logits frequency for the Performance Assessment of Self-care Skills (PASS).
The curved line represents the shape of the distribution (N = 68).

2.3.4 Descriptive functional status statistics

Table 2-16 summarizes the descriptive statistics of the functional status measures for the participants at 3 months post-stroke. For the 68 participants, the average score on the GOS5 approached 4, indicating they had moderate disabilities, and on the GOS8 approached 5, indicating they were independent in activities of daily life, but unable to resume previous activities either at work or socially. The participants demonstrated moderate to slight disabilities on the mRS, with a mean score of 2.44, indicating their abilities ranged from "unable to carry out all previous activities but able to look after own affairs without assistance" to "requiring some help but able to walk without assistance." The mean of the summative score of all BI items (in a 20-point scale) was 17.07 (equivalent to 85.37 in a 100-point scale), showing that the participants were moderately dependent on others (Anemaet, 2002; Shah, Vanclay, & Cooper, 1989). Specifically, the mean score of the PASS was 2, indicating that, on average, the participants occasionally required verbal assistance but usually not physical assistance to perform tasks (Holm & Rogers, 1999) (see Table 2-16). Overall, the global non-summative and summative ADL functional status measures revealed that stroke survivors at 3 months post-stroke may depend on assistance to perform complex ADL tasks and may be independent in the BADL tasks.

Table 2-16: Descriptive Statistics of Stroke Outcome Measures (N = 68)

Functional status measure	Type	Mean	Range of sample	SD
GOS5	Global Non-summative	3.78	3 – 5	0.86
GOS8	Global Non-summative	4.91	3 – 8	1.50
mRS	Global Non-summative	2.44	0 – 5	1.20
BI[a]	Global Summative	17.07	0 – 20	5.18
PASS[b]	Global Average	1.82	0 – 3	0.89
BI[c]	Person ability logits	2.86	-4.77 – 4.74	2.69
PASS[d]	Person ability logits	0.43	-3.79 – 4.50	1.75

Note. GOS5 = Glasgow Outcome Scale, 5-point version. GOS8 = Glasgow Outcome Scale, 8-point version. mRS = Modified Rankin Scale. BI = Barthel Index. PASS = Performance Assessment of Self-care Skills.
[a]Descriptive statistics of BI from the summative scores of all items.
[b]Descriptive statistics of PASS from the average of the independence mean scores from each task.
[c]Descriptive statistics of the BI from the Rasch person ability logits.
[d]Descriptive statistics of the PASS from the Rasch person ability logits.

2.3.5 Correlations

Table 2-17 summarizes the correlation coefficients among the measures. All the correlations were statistically significant. The relationship between the PASS and the BI was excellent (r = 0.80, $p< 0.01$). The relationships between the PASS and the GOS5, GOS8, and mRS were also excellent (r_s = -0.83 to 0.83, all $p< 0.01$). The relationships between the BI and the GOS5, GOS8, and mRS were good (r_s = -0.76 to 0.70, all $p< 0.01$). Finally, the relationships among the non-summative measures (GOS5, GOS8, and mRS) were also excellent (r_s = -0.88 to 0.95, all $p< 0.01$) (Portney & Watkins, 2000).

Table 2-17: Correlations among Stroke Outcome Measures (N = 68)

	GOS5	GOS8	mRS	BI[c]	PASS[c]
GOS5	1.00				
GOS8	0.95**[a]	1.00			
mRS	-0.84**[a]	-0.88**[a]	1.00		
BI[c]	0.63**[a]	0.70**[a]	-0.76**[a]	1.00	
PASS[c]	0.83**[a]	0.88**[a]	-0.83**[a]	0.80**[b]	1.00

Note. GOS5 = Glasgow Outcome Scale, 5-point version. GOS8 = Glasgow Outcome Scale, 8-point version. mRS = Modified Rankin Scale. BI = Barthel Index. PASS = Performance Assessment of Self-care Skills.

[a]Analyses were performed using Spearman's Rho correlations. [b] Analysis was performed using Pearson Product-Moment correlations. [c] BI and PASS data for correlation analyses were their person ability logits from the Rasch analysis.

* $p < .01$

**$p < .001$

2.3.5.1 Item Hierarchy of 5-ADL-measures combined (GOS5, GOS8, mRS, BI, and PASS)

In order to compare the five ADL measures, we used a Rasch PCM to combine the items into a single metric. Although the combined measure cannot be used clinically, it is efficient for comparing the measurement properties of the five measures, especially the three non-summative measures which are frequently used evaluation tools in clinical practice. The three non-summative measures were single scales and could not be analyzed by the Rasch analysis separately. In addition, unlike traditional statistical methods, Rasch analysis can separate out items and persons so that the properties regarding only the "item" can be confirmed.

Two hundred and eleven administrations of the 5 ADL measures (the GOS5, GOS8, mRS, BI and PASS) were analyzed by the Rasch analysis. Table 2-19 summarizes the anchored item difficulty logits of items of the 5-ADL-measures for the stroke participants. The item hierarchy was ordered based on their item difficulty logits. The GOS8 was the most difficult item in the 5-ADL-measures for the stroke participants, followed by 11 PASS task items. The 11 PASS task items were IADL tasks, except for the task of trimming toenails. These 11 tasks included some cognitive IADL (CIADL) tasks, such as financial tasks (mailing, checkbook balancing, paying bills, and shopping) and meal preparation tasks (using sharp utensils and stovetop use), and some physical IADL tasks, (changing bed linens, carrying garbage, cleanup after meal preparation, and sweeping). The following item is the mRS. The rest of the CIADL tasks in the PASS (telephone use, medication management, obtaining visual information, playing bingo, small repairs, obtaining auditory information) and two functional mobility tasks in the PASS (stair use and bathtub or shower transfers) were more difficult than the BI items. The easiest items for the stroke survivors were the 5 BI items (bowels, bladder, walking, transfer from bed to chair, and feeding) and the GOS5.

2.3.5.2 Diagnosis of construct validity: 5-ADL-measures combined (GOS5, GOS8, mRS, BI, and PASS)

5-ADL-measures – Fit statistics

Table 2-18 summarizes the anchored item difficulty logits and fit statistics of the items for the stroke participants. While some values of the outfit mean squares were out of the recommended range (6 BI items and 10 PASS task items), no item had both fit statistics values ≤ 0.5 and ≥ 1.7.

5-ADL-measures – Principal component analysis of Rasch residuals

The standard residual variance scree plot of principal component analysis (PCA) for the 5-ADL-measures is presented in Figure 2-15. The Rasch model explained 99.9% of the total variance (100%). The unexplained variance for the 5-ADL-measures was 0.1%, and the first contrast in the residuals explained less than 0.1% of the variance. The eigenvalue of the first contrast was 5.3, indicating that it had the strength of about 5 items of 38 items in this analysis. The Rasch factor analysis for the first contrast yielded factor loadings, ranging from -0.58 to 0.71 (see Table 2-19). The opposed poles of the factor loadings distinguished the common variance of the items. The "A" (checkbook balancing in the PASS) and "a" (toilet use in the BI) in the plot identified the items with the most opposed loadings on the first contrast in the residuals of the 5-tool-measures. The items with positive values of factor loadings included 11 PASS task items (checkbook balancing, bill paying, mailing, obtaining visual information, medication management, shopping, obtaining auditory information, playing bingo, home safety, telephone use and using sharp utensils) and 1 BI item (bladder), most of which were cognitive activities. In contrast, the tasks with negative values of factor loadings included the other 14

Table 2-18: Fit Statistics and Rasch Item Difficulty Logits of the 5-ADL-Measures

MOST DIFFICULT					
Instrument	Task/Item	Logits	Infit MNSQ	Outfit MNSQ	Raw
GOS8		2.31	1.02	1.01	837
PASS	Trimming toenails	1.82	1.56	2.08	239
PASS	Mailing	1.63	0.96	0.97	287
PASS	Changing bed linens	1.59	1.30	2.42	274
PASS	Checkbook balancing	1.54	1.18	1.43	293
PASS	Bill paying	1.27	1.08	0.97	331
PASS	Carrying garbage	1.24	1.39	2.18	333
PASS	Using sharp utensils	1.20	0.83	0.51	338
PASS	Stovetop use	1.13	0.63	0.36	352
PASS	Cleanup after meal preparation	1.05	0.59	0.29	362
PASS	Shopping	0.85	0.94	0.76	386
PASS	Sweeping	0.84	1.03	0.68	396
mRS		0.69	0.86	0.84	566
PASS	Telephone use	0.58	0.93	0.53	433
PASS	Medication management	0.41	1.19	1.19	432
PASS	Obtaining information-visual	0.36	1.57	4.81	462
PASS	Playing bingo	0.30	1.39	8.83	467
PASS	Small repairs	0.22	1.04	0.80	471
PASS	Stair use	0.20	1.14	1.76	477
PASS	Home safety	0.11	1.01	0.78	485
PASS	Bathtub/Shower transfers	0.08	1.24	1.65	472
PASS	Obtaining information-auditory	0.00	1.49	2.40	497
BI	Bathing	-0.38	0.70	0.45	160
BI	Stairs	-0.43	0.82	0.50	339
PASS	Dressing	-0.55	0.96	0.70	606
PASS	Indoor walking	-0.81	1.38	1.01	546
PASS	Toilet transfers	-0.99	0.90	0.62	548
PASS	Oral hygiene	-1.03	0.81	0.51	549
BI	Dressing	-1.06	0.66	0.48	348
BI	Grooming	-1.14	0.76	0.46	176
BI	Toilet use	-1.18	0.58	0.27	365
PASS	Bed transfers	-1.47	-0.93	0.76	556
GOS5		-1.49	0.65	0.56	596
BI	Feeding	-1.69	1.22	1.10	358
BI	Transfer (bed-chair)	-1.70	0.77	0.26	574
BI	Walking	-1.74	0.71	0.41	565
BI	Bowels	-1.96	1.05	0.93	387
BI	Bladder	-1.80	0.97	0.90	376
LEAST DIFFICULT					

Note. BI = Barthel Index. GOS5 = Glasgow Outcome Scale, 5-point version. GOS8 = Glasgow Outcome Scale, 8-point version. mRS = Modified Rankin Scale. Measure = Item difficulty logits. MNSQ = Mean Square. Raw = Sum of the scored responses to an item. PASS = Performance Assessment of Self-care Skills (PASS).

PASS task items (toilet transfers, indoor walking, bed transfers, dressing, sweeping, stair use, oral hygiene, carrying garbage, bathtub or shower transfers, changing bed linens, cleanup after meal preparation, trimming toenails, stovetop use, and small repairs), the other 9 BI items (toilet use, stairs, bathing, transfers from bed to chair, walking, dressing, feeding, grooming, and bowels), and the global non-summative measures (GOS5, GOS8, and mRS), most of which were motion-oriented activities. The factor loading against the item difficulty logits of each task in the PASS was also plotted (see Figure 2-16), in which the plot represented a random pattern. The items with negative factor loadings generally had a wider distribution of item difficulty levels than those with positive factor loadings. Because there was no specific group of clustered items at a certain difficulty level in the plot, further investigations for the variance distributions were not required. In addition, the item difficulty logits yielded from the primary Rasch analysis explained 35672.1/35710.1 units of the variance and the first contrast accounted for 5.3 units. The factor sensitivity ratio was 0.00014, indicating that after the Rasch measure was extracted, the first contrast influenced less than 1% of the unexplained variance in the 5-ADL-measures. Clinically, this impact was too small to influence how to measure or interpret the ADL functional status of stroke survivors. In summary, when the five ADL-measures were combined into a single metric, the construct of independence was considered unidimensional.

```
       STANDARDIZED RESIDUAL VARIANCE SCREE PLOT
 Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)
                                              Empirical       Modeled
 Total variance in observations     =      35710.1 100.0%         100.0%
 Variance explained by measures     =      35672.1  99.9%          99.9%
 Unexplained variance (total)       =         38.0    .1% 100.0%      .1%
 Unexplned variance in 1st contrast =          5.3    .0%  13.9%
 Unexplned variance in 2nd contrast =          2.8    .0%   7.4%
 Unexplned variance in 3rd contrast =          2.6    .0%   6.8%
 Unexplned variance in 4th contrast =          2.0    .0%   5.4%
 Unexplned variance in 5th contrast =          2.0    .0%   5.2%

      VARIANCE COMPONENT SCREE PLOT
       +--+--+--+--+--+--+--+--+--+
   100%+  T                       +
       |     M                    |
  V 63%+                          +
  A    |                          |
  R 40%+                          +
  I    |                          |
  A 25%+                          +
  N    |                          |
  C 16%+                          +
  E    |                          |
    10%+                          +
  L    |                          |
  O  6%+                          +
  G    |                          |
  |  4%+                          +
  S    |                          |
  C  3%+                          +
  A    |                          |
  L  2%+                          +
  E    |                          |
  D  1%+                          +
       |                          |
   0.5%+        U  1  2  3  4  5  +
       +--+--+--+--+--+--+--+--+--+
          TV MV UV U1 U2 U3 U4 U5
        VARIANCE COMPONENTS
```

Figure 2-15: Standard residual variance scree plot of principal component analysis (PCA) for the 5-ADL-Measures.

T, TV represents the total variance in the observations. M, MV represents variance in the observations explained by the Rasch measures. U, UV represents unexplained variance. 1, U1 represent first contrast (component) in the residuals. 2 (U2), 3 (U3), 4 (U4) and 5 (U5) represent second, third, fourth and fifth contrasts (components) in the residuals, accordingly.

Table 2-19: Matrix of Standardized Residual Contrast 1 of PCA for the 5-ADL-Measures.

	Instrument	Tasks / Items	Loading	Measure
A	PASS	Checkbook balancing	0.71	1.54
B	PASS	Bill paying	0.65	1.27
C	PASS	Mailing	0.58	1.63
D	PASS	Obtaining information-visual	0.48	0.36
E	PASS	Medication management	0.45	0.41
F	PASS	Shopping	0.42	0.85
G	PASS	Obtaining information-auditory	0.38	0.00
H	PASS	Playing bingo	0.36	0.30
I	PASS	Home safety	0.28	0.11
J	PASS	Telephone use	0.19	0.58
K	PASS	Using sharp utensils	0.04	1.20
L	BI	Bladder	0.03	-1.80
a	BI	Toilet use	-0.58	-1.18
b	BI	Stairs	-0.57	-0.43
c	BI	Bathing	-0.53	-0.38
d	BI	Transfer (bed-chair)	-0.52	-1.74
e	BI	Walking	-0.51	-1.70
f	PASS	Toilet transfers	-0.49	-0.99
g	PASS	Indoor walking	-0.46	-0.81
h	PASS	Bed transfer	-0.46	-1.47
i	BI	Dressing	-0.40	-1.06
j	PASS	Dressing	-0.40	-0.55
k	PASS	Sweeping	-0.33	0.84
l	mRS		-0.32	0.69
m	PASS	Stair use	-0.27	0.20
n	PASS	Oral hygiene	-0.27	-1.03
o	PASS	Carrying garbage	-0.22	1.24
p	PASS	Bathtub/Shower transfers	-0.22	0.08
q	PASS	Changing bed linens	-0.18	1.59
r	GOS5		-0.16	-1.49
s	GOS8		-0.10	2.31
S	PASS	Cleanup after meal preparation	-0.09	1.06
R	PASS	Trimming toenails	-0.09	1.82
Q	BI	Feeding	-0.05	-1.69
P	PASS	Stovetop use	-0.05	1.13
O	BI	Grooming	-0.05	-1.14
N	BI	Bowels	-0.02	-1.96
M	PASS	Small repairs	-0.01	0.22

Note. BI = Barthel Index. GOS5 = Glasgow Outcome Scale, 5-point version. GOS8 = Glasgow Outcome Scale, 8-point version. mRS = Modified Rankin Scale. Measure = Item difficulty logits. PASS = Performance Assessment of Self-care Skills (PASS).

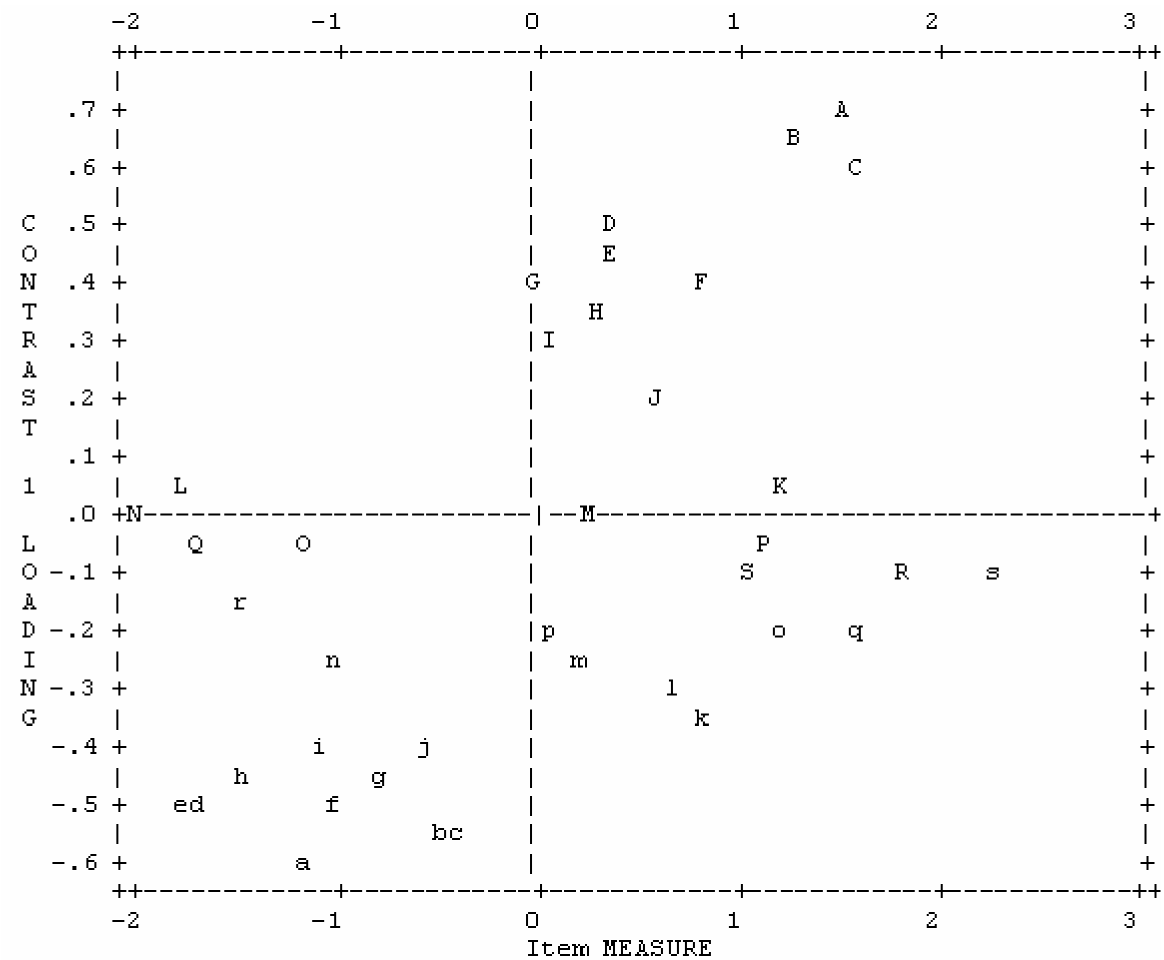

Figure 2-16: Standardized residual contrast 1 plot of principal component analysis for the 5-ADL-measures.
The plot is plotted by the contrast 1 loadings and item measures (item difficulty logits) in the matrix of standardized residual contrast 1. This plot presents the clustered residuals visually. The "A" (checkbook balancing in the PASS) and "a" (toilet use in the BI) in the plot identify the item loadings that are the most opposed on the first contrast in the residuals.

5-ADL-measures – Category function

The category function was evaluated using the partial credit model. During the analyses, each item in a measure was treated as a single scale. The data contained responses to the items in the GOS5 (scores 1-5), GOS8 (scores 1-8), mRS (score 0-5), BI (score 0-3) and the PASS (score 0-3).

Even though the scores of the global non-summative measures were recoded for the PCM analyses, Table 2-20 shows the data corresponding to the original scales. The threshold of the categories in the GOS5 showed a large gap between scores 3 and 4, indicating variances between the scores were too large. The outfit statistic for a mRS score of 5 was out of the recommended range, indicating noise in this score. In addition, some of the categories of the GOS5 and GOS8 had no responses. Thus, the category functions of these measures functioned only moderately (see Table 2-20).

The BI categories generally functioned well in this analysis. Only the score 0 for the feeding item was not within the recommended outfit range. Although some values of the thresholds did not increase with the category values, each threshold distance was within the recommended range of 1.4 to 5 logits (see Table 2-21).

In the PASS, the most problematic category was score 1 of the obtaining visual information item. Also, the category orders in the task items of bathtub or shower transfers, carrying garbage, and obtaining visual information, did not meet the expectation that more difficult items would correspond with less independent performance. Additionally, some categories in the stair use, bathtub or shower transfers, changing bed linens, obtaining auditory information, obtaining visual information, and playing bingo items, were out of the recommended range of the outfit MNSQ (i.e., outfit statistics < 2) (see Table 2-22). While these categories were not perfect in this partial credit model, the rating scale for the PASS as a whole still functioned well (see Table 2-15 and Figure 2-11). Therefore, the categories of the PASS task items were well-functioning.

Table 2-20: Category Functions for Items of the Global Non-summative Measures in Partial Credit Model of the 5-ADL-Measures

Item	Category	Observed %	OBSVD AVRGE	Structure calibration	Infit MNSQ	Outfit MNSQ
GOS5						
	1	0	N/A	N/A	N/A	N/A
	2	<1	-6.49	None	0.94	0.62
	3	42	-0.29	-6.53	0.59	0.63
	4	30	1.71	2.77	0.42	0.21
	5	27	3.08	3.75	0.78	0.75
GOS8						
	1	0	N/A	N/A	N/A	N/A
	2	0	N/A	N/A	N/A	N/A
	3	17	-1.57	None	1.34	1.26
	4	28	0.53	-2.88	0.81	0.88
	5	20	1.59	-0.94	0.72	0.65
	6	12	2.26	-0.08	0.59	0.67
	7	21	2.98	-0.30	1.37	1.44
	8	1	4.61	4.21	0.74	0.98
mRS						
	0	6	3.64	3.84	1.39	1.32
	1	21	2.97	1.66	0.72	0.73
	2	28	1.58	0.59	0.72	0.76
	3	30	0.65	-1.60	0.57	0.54
	4	11	-1.08	-4.49	0.87	0.81
	5	4	-4.41	None	3.39	2.31

Note. GOS5 = Glasgow Outcome Scale, 5-point version. GOS8 = Glasgow Outcome Scale, 8-point version. Observed % = Percentage of occurrences of that category. OBSVD AVERGE = Average of the person ability logits for all persons in the sample who had that particular response to that category. Structure calibration = The difficulties estimated for choosing one response category over another. MNSQ = Mean square. mRS = Modified Rankin Scale.

Table 2-21: Category Functions for Items of the Barthel Index (BI) in Partial Credit Model of the 5-ADL-Measures

Item	Category	Observed %	OBSVD AVRGE	Structure calibration	Infit MNSQ	Outfit MNSQ
Bowels						
	0	7	-3.98	None	0.95	0.21
	1	2	0.13	1.05	1.20	1.26
	2	91	1.60	-1.05	1.12	0.90
Bladder						
	0	6	-4.21	None	0.58	0.24
	1	9	0.19	-0.30	1.08	1.00
	2	85	1.69	0.30	1.25	1.11
Grooming						
	0	16	-1.85	N/A	0.78	0.42
	1	84	1.79	N/A	0.70	0.64
Toilet use						
	0	9	-3.37	None	0.60	0.26
	1	8	-0.24	0.19	0.67	0.21
	2	83	1.83	-0.19	0.45	0.63
Feeding						
	0	5	-3.48	None	1.74	2.17
	1	19	-0.02	-1.30	1.10	0.97
	2	76	1.83	1.30	1.15	1.08
Transfer						
	0	6	-4.15	None	1.31	1.25
	1	2	-2.13	0.27	0.24	0.03
	2	4	-0.50	0.18	0.57	0.16
	3	88	1.73	-0.45	0.75	0.81
Walking						
	0	6	-4.46	None	0.71	0.41
	1	1	-1.52	0.74	1.22	0.55
	2	12	-0.34	-1.50	0.76	0.35
	3	81	1.85	0.77	0.44	0.79
Dressing						
	0	8	-3.60	None	0.49	0.33
	1	18	0.14	-0.76	0.79	0.44
	2	74	1.98	0.76	0.64	0.77
Stairs						
	0	14	-1.99	None	1.02	0.70
	1	10	-0.20	0.41	0.66	0.28
	2	76	1.98	-0.41	0.58	0.65
Bathing						
	0	50	-1.24	N/A	-1.24	-0.83
	1	160	1.96	N/A	1.96	1.83

Note. Observed % = Percentage of occurrences of that category. OBSVD AVERGE = Average of the person ability logits for all persons in the sample who had that particular response to that category. Structure calibration = The difficulties estimated for choosing one response category over another. MNSQ = Mean square.

Table 2-22: Category Functions for Items of the Performance Assessment of Self-Care Skills (PASS) in Partial Credit Model of the 5-ADL-Measures

Item	Category	Observed %	OBSVD AVRGE	Structure calibration	Infit MNSQ	Outfit MNSQ
Bed transfers						
	0	6	-4.23	None	1.21	0.98
	1	4	-1.48	-0.72	0.82	0.53
	2	10	0.54	0.33	0.96	0.81
	3	80	1.80	0.39	0.80	0.90
Stair use						
	0	20	-1.46	None	1.05	3.65
	1	4	0.60	1.11	1.56	0.81
	2	3	1.01	0.89	0.64	0.20
	3	72	1.99	-2.00	1.19	1.05
Toilet transfers						
	0	9	-3.47	None	0.96	1.00
	1	2	-0.59	0.50	1.04	0.40
	2	9	0.40	-0.36	0.81	0.57
	3	80	1.83	-0.14	0.81	0.84
Oral hygiene						
	0	9	-3.45	None	0.94	0.84
	1	1	-1.24	0.99	0.64	0.15
	2	10	0.24	-1.03	0.83	0.46
	3	80	0.86	0.04	0.69	0.81
Bathtub/Shower transfers						
	0	15	-2.16	None	0.64	0.52
	1	11	1.06	-0.21	1.77	3.38
	2	8	1.05*	0.96	1.04	0.31
	3	66	2.01	-0.75	1.47	1.37
Trimming toenails						
	0	61	0.26	None	1.73	3.21
	1	1	1.52	3.29	1.52	0.10
	2	1	1.80	-0.57	1.62	0.27
	3	37	2.72	-2.71	1.43	2.22
Dressing						
	0	9	-3.32	None	0.74	0.58
	1	5	-0.05	-0.29	1.09	0.67
	2	21	0.73	-0.54	0.89	0.60
	3	64	2.09	0.83	1.10	1.00
Shopping						
	0	22	-1.41	None	0.78	0.71
	1	7	0.62	0.34	0.64	0.33
	2	37	1.58	-1.53	0.66	0.56
	3	34	2.55	1.19	1.45	1.25
Bill paying						
	0	35	-0.55	None	1.05	0.91
	1	8	1.01	0.63	0.79	0.33
	2	22	1.60	-0.99	1.08	0.90
	3	35	2.71	0.35	1.16	1.50
Checkbook balancing						
	0	44	-0.21	None	1.03	0.90
	1	4	1.59	1.45	1.66	0.87
	2	19	1.70	-1.60	1.17	1.01
	3	32	2.76	0.16	1.33	2.35

Table 2-23 (Continued)

Mailing						
	0	43	-0.30	None	0.81	0.68
	1	4	1.23	1.32	1.32	0.54
	2	26	1.92	-1.98	1.00	1.29
	3	27	2.88	0.67	1.11	1.05
Carrying garbage						
	0	45	-0.16	None	1.35	3.69
	1	<1	1.83	3.90	2.76	0.17
	2	5	1.19*	-2.27	1.43	1.19
	3	50	2.43	-1.63	1.41	1.47
Telephone use						
	0	28	-0.99	None	0.99	0.66
	1	1	-0.15	2.46	0.34	0.10
	2	8	1.02	-1.43	1.03	0.24
	3	63	2.21	-1.03	0.85	0.88
Medication management						
	0	17	-1.64	None	1.29	1.44
	1	5	-0.24	0.48	0.70	0.99
	2	35	1.33	-1.58	1.10	1.14
	3	44	2.32	1.10	1.27	1.13
Changing bed linens						
	0	52	0.04	None	1.29	2.26
	1	5	1.49	1.51	1.83	4.44
	2	2	1.75	0.93	0.29	0.03
	3	40	2.63	-2.44	1.30	1.96
Obtaining information-auditory						
	0	18	-1.47	None	1.60	3.53
	1	3	-0.56	1.35	0.90	0.91
	2	5	0.75	0.05	0.51	0.60
	3	75	1.92	-1.40	1.65	5.73
Obtaining information-visual						
	0	24	-0.98	None	1.49	2.01
	1	1	-1.69*	2.75	0.73	5.30
	2	7	1.54	-1.55	1.60	6.17
	3	69	1.96	-1.20	1.74	8.29
Small repairs						
	0	20	-1.55	None	1.13	1.41
	1	1	-0.47	2.49	0.99	0.57
	2	15	0.90	-2.31	0.89	0.37
	3	64	2.12	-0.18	0.97	0.86
Sweeping						
	0	34	-0.66	None	1.02	0.75
	1	1	0.48	3.04	0.10	0.01
	2	7	1.36	-1.76	0.76	0.47
	3	58	2.28	-1.28	1.10	0.96
Indoor walking						
	0	10	-3.08	None	1.43	1.19
	1	5	-0.24	0.04	1.23	0.05
	2	2	0.96	1.81	0.94	1.68
	3	84	1.77	-1.86	1.57	1.15

Table 2-23 (Continued)

Home safety						
	0	19	-1.75	None	0.91	0.60
	1	2	0.05	1.78	0.38	0.05
	2	10	1.44	-1.10	1.16	1.05
	3	70	1.97	-0.69	1.25	1.04
Playing bingo						
	0	22	-1.08	None	1.45	9.90
	1	3	-0.28	1.60	0.85	1.91
	2	5	1.10	-0.08	1.21	1.24
	3	70	1.99	-1.52	1.42	6.13
Stovetop use						
	0	40	-0.52	None	0.61	0.44
	1	1	0.59	2.66	0.83	0.13
	2	11	1.52	-1.97	0.35	0.11
	3	48	2.55	0.69	0.72	0.57
Using sharp utensils						
	0	43	-0.34	None	0.84	0.54
	1	3	0.81	2.06	0.93	0.70
	2	5	1.22	-0.46	1.10	0.21
	3	50	2.54	-1.60	0.78	0.57
Cleanup after meal preparation						
	0	40	-0.51	None	0.62	0.37
	1	2	0.77	2.45	0.53	0.06
	2	4	1.33	-0.59	0.25	0.03
	3	54	2.48	-1.86	0.59	0.42

Note. Observed % = Percentage of occurrences of that category. OBSVD AVERGE = Average of the person ability logits for all persons in the sample who had that particular response to that category. Structure calibration = The difficulties estimated for choosing one response category over another. MNSQ = Mean square.

5-ADL-measures – Scale linearity

The scale linearity of the PASS is shown in Figure 2-17. The raw data were the sum of the scored responses to each item (see Table 2-18) and the measures were the item difficulty logits of each item. The R^2 linearity value is 0.03, indicating that about 3% of the variance can be explained by the model (Portney & Watkins, 2000). It also indicates that the scale of the 5-ADL-measure was not linear.

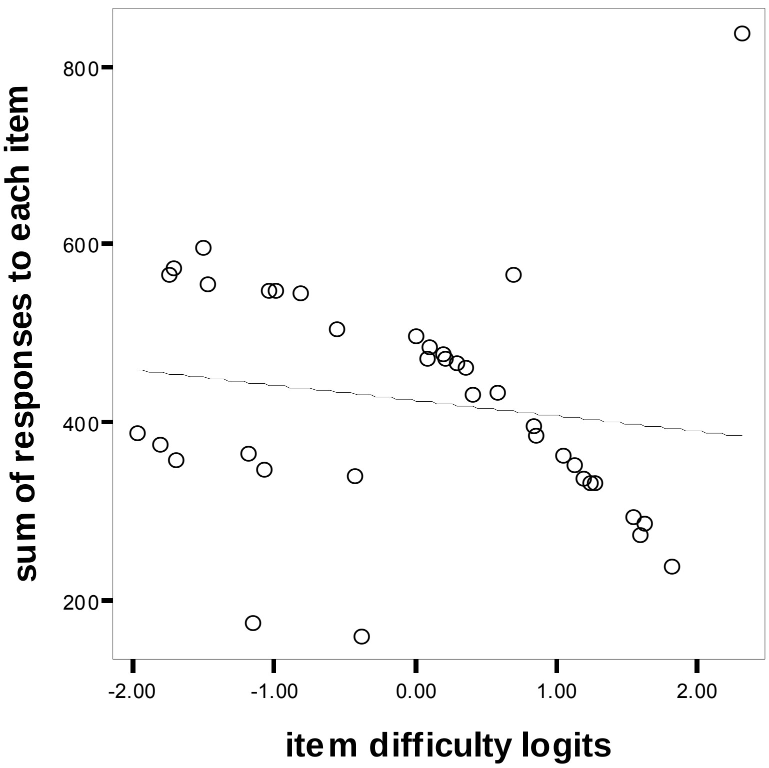

Figure 2-17: Scale linearity of the 5-ADL-measures.
Each dot represents one item of the 5 ADL measures. They were plotted by their item difficulty logits (x-axis) and sum of responses to each item in the Rasch analysis (y-axis). The line was the ordinary least squares regression of y on x.

2.3.5.3 Diagnosis of reliability: 5-ADL-measures combined (GOS5, GOS8, mRS, BI, and PASS)

5-ADL-measures – Item reliability

The item reliability index of the 5-ADL-measures was 0.99, which indicated that the item difficulty hierarchy of the 5-ADL-measures would be consistent over other populations with the same traits and same sample size as our data.

5-ADL-measures – Person reliability

The person reliability index of the 5-ADL-measures was 0.91, indicating a good to excellent reproducibility. It also implied that the persons who were more independent could be reliably distinguished from those who were less independent by the 5-ADL-measures. The person separation index of the 5-ADL-measures was 3.28, which indicated an excellent stability of the person ordering. The internal consistency of the 5-ADL-measures was also excellent with a Cronbach's alpha of 0.97. The above data showed that the 5-ADL-measures are extremely reliable to evaluate ADL performance for stroke survivors.

2.3.5.4 Person ability logits of the 5-ADL-measures combined (GOS5, GOS8, mRS, BI, and PASS)

The Rasch person ability logits of the 5-ADL-measures were calculated based on the anchored item values of this tool. These Rasch person ability logits represented the ADL performance of the participants at 3 months post-stroke in the 5-ADL-measures. Figure 2-18 presents the item-person map for the 5-ADL-measures. The map illustrates the distributions of the abilities of the persons at 3 months stroke, along the ability logit scale, and the difficulties of the items from the 5 measures, along the difficulty logit scale, at the same time. The distribution of the persons was as clustered as for the items. Figure 2-19 was further plotted to show the person ability logits distribution of the 5-ADL-measures. The curve in the graph was normally distributed, implying that the items of the 5 ADL measures were responsive to different levels of functional status of the stroke survivors at 3 months post-stroke. Furthermore, the item-person map of the 5-ADL-measures showed that a stroke survivor whose ability logit value was the same as the difficulty logit value of the GOS8 had a 50 % probability of being rated the highest score (full recovery) in the GOS8, and would have a greater chances of being rated fully independent in all BADL and IADL tasks in the PASS, BI, mRS and GOS5, which are located lower on the item difficulty hierarchy.

```
    Persons MAP OF Items
       <more>|<rare>
 5            +
              |
              |
         XX   |
            T|
 4            +
              |
         XX   |
              |
          X   |
 3     XXXX   +
          X   |
         XX S|
         XX   |T GOS8
          X   |
 2            +
        XXX   |  P_teonail
        XXX   |  P_linens      P_check      P_mail
   XXXXXXXX   |
     XXXXXX   |S P_bend        P_stovetop   P_utensil    P_bill
 1           M+  P_clean
       XXXX   |  P_sweep       P_shop
       XXXX   |  P_tel         mRS
      XXXXX   |  P_med         P_vision     P_bingo
        XXX   |  P_stair       p_repair     P_safety
 0     XXXX   +M P_tub         P_aud
          X   |
        XXX S|   BI_Bathing    BI _Stairs
          X   |  P_dress
              |  P_walk
-1        X   +  P_toilet      BI_Dressing    P_oral
          X   |S BI_Grooming    BI_Toilet Use
              |  P_bed         GOS5
              |  BI_Feeding    BI_Transfer
          X   |  BI_Bladder    BI_Walking
-2           T+  BI_Bowels
              |
              |T
              |
          X   |
-3            +
              |
              |
              |
              |
-4            +
              |
              |
          X   |
              |
-5      XXX   +
       <less>|<frequ>
```

Figure 2-18: Item-person map for the5-ADL-measures.

To the left of the dashed line (the logit scale) is the person ability scale, distributed from the greatest ability at the top to the least ability at the bottom. Each X represents a person in the sample. To the right of the dashed line is the item difficulty scale distributed from the most difficult item at the top to the easiest item at the bottom. Item names with "P_" in the front represent the items of the PASS. Items names with "BI_" in the front represent the items of the BI. Along the logit scale, "M" represents the mean person ability or item difficulty estimate, "S" represents the location of one standard deviation (SD) from the mean estimates, and "T" is the second SD away from the mean estimates.

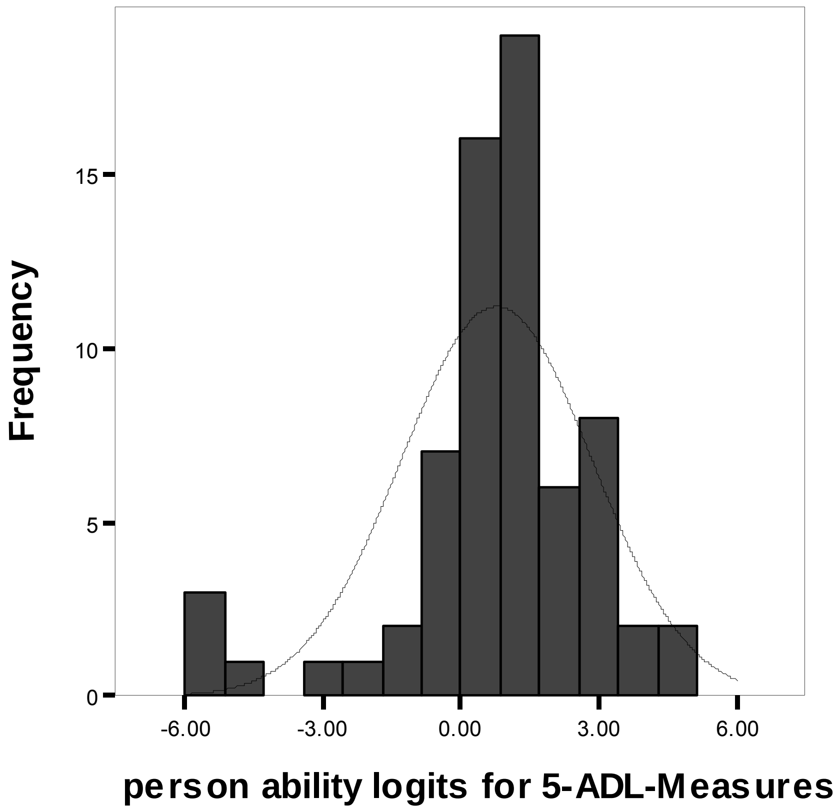

Figure 2-19: Histogram of the person ability logits frequency for the 5-ADL-measures. The curved line represents the shape of the distribution (N = 68).

2.4 DISCUSSION

The purpose of current study was to confirm, correlate, and compare five ADL measures, which used different scoring systems to measure the functional status of stroke survivors, at 3 months post-stroke. The separate validities and reliabilities of each tool were first confirmed from the literature, and then Rasch analysis was used to convert BI and the PASS ordinal scores to interval scores so that the relationships among the tools could be established. Because each of the five ADL tools purported to measure independence, we then combined the items using the Rasch partial credit model (PCM), so that we could compare them on a common metric. Our findings indicated that our hypotheses were partially supported, and that more than one of the 5-ADL combined measure diagnostics, the scale linearity and category function analyses, indicated that the five ADL tools are not interchangeable nor do they provide the same information about independence in stroke survivors at 3 months post-stroke.

Our hypothesis that the relationship between the rehabilitation-relevant BI and the PASS measures would be stronger than between the BI and the global non-summative measures (GOS5, GOS8, and mRS) was supported. However, our hypothesis that the relationship between the BI and the PASS would be stronger than the relationship between the PASS and the global non-summative measures was not supported. Overall, relationships among the measures revealed that stroke survivors who performed well on one measure also performed well on the others. Although the strong relationships among the three global non-summative measures were expected (Tilley et al., 1996), and the relationships between the non-summative measures and the BI and the PASS were moderate to strong, these tools are not necessarily interchangeable. That is because, in addition to independence in daily activities, content in the GOS5, GOS8, or mRS also included level of consciousness, signs/symptoms, recovery, complaints, and level of nursing care (Dittmar, 1997).

Rasch analysis diagnostics of the BI, PASS, and the 5-ADL-combined measure indicated that they were all unidimensional -- independence was the main construct being measures. The 5-ADL combined measure was the most unidimensional, with a 99.9% explained variance in the model and the fitted items. It indicated that the items from the five instruments as a whole measured a similar construct, that is, independence. However, despite the fact that the unidimensionality was excellent, the independence was measured from multiple perspectives. The three global non-summative measures included not only independence, but also additional concepts/constructs (e.g., symptoms, recovery, complaints, level of consciousness). Because of this, the category functions of the three non-summative measures (the GOS5, GOS8, and mRS) were problematic, and the scale of the combined measure was no longer linear, as it had been for the PASS, and less so for the BI. The consequence of the non-linear 5-ADL-measure scale was that the combined tool could not sensitively detect small changes in the independence of stroke survivors and therefore would not be useful clinically (Portney & Watkins, 2000). The non-linearity of the scale also suggests that the tools are not interchangeable.

The mixing of concepts/constructs was also seen in the item-person map. Even though the item hierarchy and person order were stable for this metric, the item-person map of the 5-ADL-measures showed that the rating scores of the three non-summative measures produced contradictory interpretations. For example, the three non-summative measures did not cluster together, but were distributed over the total hierarchy, with the GOS8 being the hardest, the mRS in the middle, and the GOS5 being the easiest. Also, persons whose abilities were at the same level as the GOS5 or mRS, had a 50% probability of being rated with the highest score on the GOS5 (i.e., good recovery) or the mRS (i.e., no symptoms at all), but the map shows that for the GOS5, those participants would only have reported being able to accomplish the most basic BADL items on the BI, and only 1 of the PASS items (bed transfers). Likewise for the mRS, the map indicates that those participants would still have problems with the many CIADL. Therefore, the item-map of the 5-ADL-measures shows that the multiple attributes in the content of the global non-summative scales impacts their interpretation of independence in stroke survivors at 3 months post-stroke.

Our data at 3 months post-stroke, also showed that most participants were able to take care of their BADL, which was supported by the literature that indicated early mobilization and return to self-care are usually the primary intervention goals in acute stroke rehabilitation (Tyson & Turner, 2000; Woodson, 2002). At 3 months post-stroke and post acute care, our participants were still challenged in the performance of IADL. Moreover, when participants self-reported their functional status on the BI, they tended to overestimate their abilities. This was apparent on the item-person map with items that were common to both the BI and the performance-based PASS. Of the 7 common items, all PASS items were mapped as more difficult on the item hierarchy than their BI counterparts. The different assessment methods could be responsible for the difference. Findings from previous literature with other populations has shown that performance-based observational (PBO) assessments allow researchers or rehabilitation practitioners to gather more detailed and specific information about the processes of executing ADL tasks (Finlayson, Havens, Holm, & Denend, 2003), and that hey tend to be more valid and reliable than informant-based assessments (Rogers & Holm, 1998; Sinoff & Ore, 1997).

Limitations and Recommendations

The current study had some limitations. Our sample included participants with ischemic and hemorrhagic strokes, and the different mechanisms in the two types of stroke may have differentially influenced the severity of disabilities in the sample (Woodson, 2002). Also, because this was a secondary analysis of a completed dataset, some desirable variables were not available in the original dataset, such as specific brain lesion locations. In addition, all the rich information yielded from the Rasch analysis could not be addressed in this study because the current study was delimited to exploration of the construct of independence among the five ADL measures. Future studies should use a more homogeneous sample (e.g., ischemic stroke) to explore global non-summative, global summative, and task-specific measures that are commonly used to determine the functional status of stroke survivors beyond discharge from rehabilitation. Method of assessment should also be considered in such analysis (e.g., clinical observation, self-report, performance-based observation).

2.5 CONCLUSION

The current study compared the constructs of four common ADL functional status measures (mRS, GOS5, GOS8, BI) used with stroke survivors and one task-specific, performance-based observational assessment (PASS). Although the five measures were significantly and moderately to strongly correlated, the tools behaved differently when combined in a common metric. Rasch analysis showed that individually the PASS and the BI had excellent psychometric properties, and were valid and reliable. When the five ADL tools were combined into a single metric, the combined tool was unidimensional, but the three global non-summative measures (mRS, GOS5, GOS8) made the combined measure non-linear, because in addition to independence, the three tools also measured other concepts/constructs. Clinically, the data from the combined measure should be interpreted carefully, although it does show clearly that stroke survivors at 3 months post-stroke overestimate their abilities with self-report, when compared to the results of a performance-based measure.

3.0 COMPARISON OF STROKE SURVIVORS' SPECIFIC DISABILITY ITEMS FROM TWO ADL MEASURES: THE PERFORMANCE ASSESSMENT OF SELF-CARE SKILLS AND BARTHEL INDEX

3.1 BACKGROUND AND SIGNIFICANCE

Stroke is prevalent and disability associated with stroke is significant (American Heart Association, 2007). The disabilities related to stroke include functional limitations in the "capacity to carry out any activity within the normal range achievable by a human being," (Carod-Artal, Gonalez-Gutierrez, Herrero, Horan, & de Seijas, 2002, p. 207). Clinically, a functional status assessment is used to identify disabilities in activities of daily living (ADL) such as eating, dressing, and managing financial tasks. Functional status assessments allow rehabilitation practitioners to document patients performing everyday activities to "help to explain, confirm, or cast doubt" (Rubenstein, Calkins, Greenfield, et al., 1988, p. 563) on their functional status. However, most ADL measures utilize ordinal scales, which usually categorize human behaviors that have multiple attributes (Merbitz et al., 1989; Kenneth J. Ottenbacher, 1997; Rogers & Holm, 1998). The results from these measures can only illustrate global phenomena of ADL performance. Because the measurement properties of the ADL measures had been examined (see details in the second chapter), the present study aimed to compare specific disability items for stroke survivors at 3 months post-stroke on two rehabilitation-relevant measures, the Performance Assessment of Self-Care Skills (PASS) and the Barthel Index (BI). A second aim of the study was to identify a hierarchy of item difficulty for stroke survivors at 3 months post-stroke that could provide guidance regarding the preferred sequence of ADL items for intervention.

3.1.1 Stroke survivors and Functional Status

Understanding and interpreting ADL functional status in stroke survivors needs to be precise and comprehensive. The interpretation of functional status data significantly influences the rehabilitation practitioner's understanding of the stroke survivor's disability and affects the clinical reasoning that is used to develop treatment plans (Dittmar, 1997; Rogers, 1983). ADL data provide the most concrete and detailed information about disabilities of stroke survivors (van Boxel et al., 1995). Duncan, Jorgensen, and Wade (2000) also suggested that ADL should be the primary outcome measure in stroke due to its relative objectivity, simplicity, and relevance to patients. Moreover, improved functional status contributes to patient satisfaction with treatment as well as minimizing the potential cost of long-term care (Pamela W. Duncan et al., 2005).

3.1.2 Domains and Assessment Methods of Functional Status Measure in Stroke

In functional status measures, ADL tasks are classified into two fundamental categories: basic and instrumental ADLs (BADL and IADL) (Reuben & Solomon, 1989). BADL refers to self-maintenance activities involving functional mobility and personal care (Rogers & Holm, 1998). Sample BADL tasks are ambulation or wheelchair mobility, transfers, feeding, hygiene, toileting, and bathing (Rogers & Holm, 1998). The more complex IADL often refer to home management tasks and tasks required for independent living in the community. Sample IADL tasks are shopping, cooking, housekeeping, laundry, use of transportation, managing money, managing medicine, and use of the telephone (Lawton & Brody, 1969). BADL can be further delineated into domains of functional mobility and personal care and IADL domains with a cognitive or physical emphasis (Holm & Rogers, 1999). Assessment methods available to evaluate ADL functions include both informant-based and performance-based observational (PBO) assessments (Rogers & Holm, 1998). Informant-based assessments apply the method of asking questions via interviews or questionnaires, whereas PBO assessments use the approach of observing clients' ADL performance in natural or laboratory conditions. PBO assessments tend to be more valid and reliable than informant-based assessments (Rogers & Holm, 1998; Sinoff & Ore, 1997). Also, PBO assessments allow researchers or rehabilitation practitioners to gather more detailed and specific information about the processes of executing ADL tasks (Finlayson et al., 2003).

3.1.3 Global versus Task Specific Scoring Systems

In addition to the consideration of assessment methods, functional status scoring systems can be global or task specific (see Table 3-1). Summative global ADL measures yield summary scores by simply adding up the obtained ratings on all items to indicate severity of disability, or using a grand mean of like items to indicate severity of disability for specific ADL domains (Holm & Rogers, 1999). Examples of measures using a summative global scoring system are the Barthel Index (BI) and the Functional Independence Measure (FIM) (Ravaud et al., 1999; Ring et al., 1997; van Hartingsveld et al., 2006). Based on summative global scores, rehabilitation practitioners can only obtain a broad impression of the patient's performance (i.e., a higher sum score on the BI indicates "more independence in BADL"). Another limitation of a global score is that the exact functional ability of patients cannot be determined, thus decreasing these measures' usefulness (van Hartingsveld et al., 2006). In the FIM, the ADL data consist of summary scores of motor and cognitive items (i.e., a higher sum score from the FIM motor items indicates "more independence in functional mobility and personal self-care" and a higher sum score from cognitive items indicates "more independence in communication and cognitive function.") (Ring et al., 1997). However, use of the summary FIM scores to compare or describe functional status should be carefully considered, because they refer to a heterogeneous content (Ravaud et al., 1999). Also, global scoring on the FIM cognitive items (e.g., an item score of 4 refers to "minimal assistance with clients exerting 75% plus effort) has led to lower inter-rater reliability (Dallmeijer et al., 2005; K. J. Ottenbacher, Hsu, Granger, & Fiedler, 1996). Additionally, a potential inherent problem with the 7-category rating scale that affects the internal construct validity of the FIM has also been reported (Lundgren-Nilsson et al., 2005; K. J. Ottenbacher et al., 1996; Ring et al., 1997). Lundgren-Nilsson et al. (2005) found that, based on Rasch analysis, a 3-5 category rating scale for FIM motor items would be more appropriate than a 7 category scale to detect the level of needed assistance and to give optimal discrimination.

Using a global average score derived from individual task ratings is another manner of global scoring. In contrast to a summative global measure, this method provides information about disability specific to ADL tasks (Rogers et al., 2003). Although still global in nature, it provides greater direction to rehabilitation practitioners, because an ADL task is composed of critical subtasks and they are clinically relevant. An example of a functional status measure that uses ADL domain averages is the Performance Assessment of Self-Care Skills (PASS), which is also a criterion-referenced, performance-based observational tool (Holm & Rogers, 1999).

In contrast to a global scoring system, a task specific scoring system is derived from the ratings of single tasks with critical and essential standards (i.e., criterion-referenced tests). An example of a measure using a specific scoring system is the PASS. In the PASS, each task contains subtasks that are sequenced and performance criteria that are to be rated. For independence, the rating is based on the level of assistance provided for each subtask (Holm & Rogers, 1999). For example, the PASS item "Stovetop Use" consists of 11 subtasks, with clear, observable criteria to be rated (e.g., Turns burner off promptly [+/-- 1 minute of removing soup from burner]). The data from a specific scoring system result in better reliability between raters because of clear criteria for rating performance (Rogers et al., 2001), and they provide a meaningful functional status for rehabilitation practitioners to identify to what degree the patient can perform a specific daily task independently (Rogers, Holm, & Stone, 1997) (see Table 3-2).

Table 3-1: Examples of Global Functional Status (Summative and Average) Scoring Systems

<table>
<tr><th rowspan="2">Functional Status Measures</th><th rowspan="2">Type</th><th colspan="2">Sample Scoring</th></tr>
<tr><th>Score</th><th>Description</th></tr>
<tr><td rowspan="3">Barthel Index Items
(20 pt. version)

Bowels, Bladder, Grooming, Toilet use, Feeding, Transfer (bed-chair), Walking, Dressing, Stairs, Bathing</td><td rowspan="3">Global Summative</td><td colspan="2">Higher scores indicate greater independence in activities of daily living. Scores are derived by adding up item scores.
Sample item: BI10 - Bathing</td></tr>
<tr><td>1</td><td>Independent (May use bathtub, shower or sponge bath; Subject must be able to perform all functions without another person being present</td></tr>
<tr><td>0</td><td>Dependent (Cannot meet criteria)</td></tr>
<tr><td rowspan="3">Performance Assessment of Self-Care Skills (PASS) Task Items

Bed transfers, Stair use, Toilet transfers, Oral hygiene, Bathtub/Shower transfers, Trimming toenails, Dressing, Shopping, Bill paying, Checkbook balancing, Mailing, Carrying garbage, Telephone use, Medication management, Changing bed linens, Obtaining information-auditory, Obtaining information-visual, Small repairs, Sweeping, Indoor walking, Home safety, Playing bingo, Stovetop use, Using sharp utensils, Cleanup after meal preparation</td><td rowspan="3">Global Average [Task-specific]</td><td colspan="2">Higher scores indicate greater independence in activities of daily living. Domain scores are derived by calculating a grand mean from Domain item means.</td></tr>
<tr><td>3</td><td>Independent (No assists given for task initiation, continuation, or completion)</td></tr>
<tr><td>0</td><td>Dependent (Total assistance given, or continuous physical guidance or physical support during tasks)</td></tr>
</table>

Table 3-2: Example of Performance of Self-Care Skills (PASS) Subtask Criteria

Sample task: Stovetop use item	
1	Open soup can correctly (cut is even, entire top is off or < ½ is retained in one place)
2	Removes/handles soup can lid correctly (lifts lid with knife; punches lid into can; does not cut finger)
3	Pours/spoons soup into pan without spilling (no soup on patient, counter, or floor)
4	Adds water correctly (add 1 can of water; does not spill on self, floor)
5	Places pan on correct stove burner (burner closest to pan size)
6	Turns burner on correctly (manipulates knob for burner that soup is on or is placed onto later; sets control on medium to high)
7	Monitors soup adequately (stirs; alters heat as necessary, soup does not stick on pan, checks to make sure soup temperature is hot rather than lukewarm to touch or taste or that soup boils/bubbles)
8	Removes pan from burner when soup is still hot (steam can be seen rising from pan; checks to make sure soup temperature is hot rather than lukewarm to touch or taste)
9	Turns burner off promptly (+/- 1 minute of removing soup from burner)
10	Transports & pours soup into bowls correctly (uses mitt under pan or slides pan across counter for stability if weakness or tremor present; does not spill on floor; only minor drips on counter)
11	Transports bowls to table correctly (uses mitt under bowl or uses cart if weakness, tremor or instability present; uses bowl rim to carry; does not spill on floor)

3.1.4 Utility of Global versus Specific Functional Status Measures

Although several studies reported functional status data at 3 months post-stroke, most reported data from informant-based and global ADL measures (e.g., the FIM and the BI) (Caelo et al., 2003; Pamela W. Duncan, Bode, Lai, & Perera, 2003; Lai & Duncan, 2001; Lyden et al., 1999; Nuutinen et al., 2006). In addition to the issues of assessment methods and scoring systems, the usefulness of the scores from global ADL measures for rehabilitation practitioners is limited (Lindeboom et al., 2003). That is, these measures do not provide enough detail to guide intervention. In rehabilitation, it is extremely important to have an understandable measure that both describes the functional status of a patient and guides intervention (Smith & Taylor, 2004). The consequence of no specificity for BADL and IADL task performance in global ADL measures is that there is not enough information to guide the clinical reasoning of rehabilitation practitioners for developing suitable interventions. No evidence could be found that compared global ADL stroke functional status measures and specific PBO ADL stroke functional status measures.

3.1.5 Stroke Functional Status Hierarchies and Rasch Analysis

Linedeboom, Vermeulen, Holman, and Haan (2003) mentioned that traditional statistical methods that assume that data are on an interval scale cannot solve the difficulties related to ordinal scale scores for ADL assessments. Owing to the ordinal structure of the item set, the scores at the same scale levels may be different because the item distances along the scales are not even and because they may represent varied meanings (Merbitz et al., 1989) (for examples see Table 3-1). Furthermore, the major limitation of the classical statistical methods "is that the examinee characteristics and test characteristics cannot be separated" (Hambleton, Swaminathan, & Rogers, 1991, p. 2). This conceptual problem influences the interpretation of results, and leads to difficulty when comparing the performance of different populations or the performance between different tests. Rasch analysis is developed to reduce such limitations. It transforms ordinal data into interval data, called logits (Hambleton et al., 1991; Tesio, 2003; Wright & Linacre, 1989). This method brings an advantage, namely, "to use a single synthetic index to describe the full spectrum of functioning" (Barberger-Gateau, Rainville, Letenneur, & Dartigues, 2000, p. 310). Rasch analysis also yields two parameters– item difficulty and person ability – that allow the item and person information to be interpreted at the same time by using the same measurement units. In this way, the hierarchical pattern of ADL performance in stroke survivors can be revealed, and so can each person's ability.

Several studies have investigated the hierarchical pattern of ADL performance in stroke survivors by using Rasch analysis (Pamela W. Duncan et al., 2003; Granger, Dewis, Peters, Sherwood, & Barrett, 1979; Hsueh, Wang, Sheu, & Hsieh, 2004; Lundgren-Nilsson et al., 2005; van Boxel et al., 1995). Among these studies, however, most covered only a narrow range of daily tasks. Only Hsueh et al. (2003) and Duncan et al. (2003) evaluated both BADL and IADL tasks. Additionally, the assessment methods were informant-based (i.e., telephone or face-to-face interview) and their studies were designed to validate the tools they were developing. The evidence of a hierarchical pattern of performance in a wide range of ADL tasks in stroke survivors is still immature, and not known for patients at 3 months post-stroke.

3.1.6 Hypothesis

The present study aimed to compare specific disability items for two ADL measures used with stroke populations, the PASS (task-specific measure) and the BI (global measure), for stroke survivors at 3 months post-stroke. The items from the two measures would be converted to a single metric, using Rasch analysis, and then be compared. We also hypothesized that overall task difficulty would be greater for PASS task items than for BI items, because the PASS includes 13 IADL items, and the BI focuses only on BADL.

3.2 METHODS

3.2.1 Participants

The participants for the study were recruited in a large prospective stroke outcome study: The Use of Radiological Data to Describe, Differentiate, and Predict Impairment, Disability, and Quality of Life in Stroke/ Transient Ischemic Attack Survivors (IRB# 010593) The inclusion criteria were (a) admission to the University of Pittsburgh Medical Center (UPMC) Health system (b) diagnosis of acute stroke or a transient ischemic attack (c) radiological data available from admission assessments (e.g., CT scan, MRI, and or Xenon quantitative cerebral blood flow [ExCT qCBF]) or pharmacological interventions (tPA), and (d) physician approval. Because the pharmacological intervention, tissue plasminogen activator (tPA), used in the previous study was not approved for use in children, children were excluded from the study. No exclusion criteria based on gender, race, ethnicity, or HIV status were applied.

3.2.2 Procedures

In the large prospective stroke outcome study, the study team requested informed consent directly from the patient or the appropriate proxy within 24 hours after admission to UPMC Presbyterian Hospital with a primary diagnosis of acute stroke. The attending physician or nursing staff gathered and documented demographic data after eligibility requirements were verified. The ADL assessments were administered by the staff of the Department of Occupational Therapy at 24 hours, 5 days, and 3, 6, 9, and 12-months post acute stroke event. The staff were trained by the PI in the assessment procedures to a minimum interobserver standard of >90% with the criterion assessor. If participants lived within a 150 mile radius of Pittsburgh, the ADL assessments took place where the person was residing at the time. The University of Pittsburgh Institutional Review Board (IRB#010593) approved the study and recruitment procedures also followed the guidelines of confidentiality under HIPAA.

3.2.3 Instruments

ADL functional status of the patients at 3, 6, 9, and 12-months post-stroke was measured by global and criterion-referenced specific ADL measures. The global ADL measure was the Barthel Index (BI) (Wade & Collin, 1988). The criterion-referenced specific ADL measure was the Performance Assessment of Self-Care Skills (PASS) (Holm & Rogers, 1999).

3.2.3.1 Barthel Index (BI)

The BI is a frequently used instrument in stroke clinical research. It measures independence in 10 basic activities of daily living (BADL) (bowel and bladder control, toileting, transfers, wheelchair mobility, stairs, grooming, bathing, feeding, and dressing). The BI has high interrater reliability ($r_s > .89$, $p < .001$) and correlates highly with performance-based functional measures ($r_s > .80$, $p < .01$) (Shinar et al., 1987). The construct validity of the BI was confirmed by factor analysis, indicating the BI items measure the same domain. The BI was also shown to measure a similar domain of ADL as other ADL measures such as Katz Index of ADL (concurrent validity) (Wade & Collin, 1988). The predictive validity of the BI was confirmed in a study, in which the stroke survivors who had higher BI scores at discharge had better outcomes at 6 months after discharge (Granger et al., 1988). Each item has its own rating criteria and scoring ranges (from 0 to 3) (see Table 3-3). Generally, lower scores indicate less independence in performing the activity, whereas higher scores indicate more independence. In this study, the BI was administered in a self-report interview format.

Table 3-3: Barthel Index Items and Rating Scale

Items	Score	Criteria
Bowels		
	2	Continent
	1	Occasional accident
	0	Incontinent
Bladder		
	2	Continent
	1	Occasional accident
	0	Incontinent / catheterized and unable to manage
Grooming		
	1	Independent for face/hair/teeth/shaving
	0	Needs help
Toilet use		
	2	Independent
	1	Needs some help
	0	Dependent
Feeding		
	2	Independent in all actions
	1	Needs help (e.g., cutting, spreading butter)
	0	Dependent
Transfer (bed-chair)		
	3	Independent
	2	Minor help (verbal or physical)
	1	Major help, can sit
	0	Unable
Walking		
	3	Independent (may use aid)
	2	Walks with the help of a person (verbal/physical)
	1	Independent in wheelchair
	0	Unable
Dressing		
	2	Independent (including buttons/zips/laces)
	1	Needs help, but does half
	0	Dependent
Stairs		
	2	Independent
	1	Needs help (verbal/physical)
	0	Unable
Bathing		
	1	Independent
	0	Dependent

3.2.3.2 Performance Assessment of Self-Care Skills (PASS)

The PASS is a criterion-referenced and performance-based instrument used to rate the independence, safety and adequacy of task performance. It contains of 26 tasks and 163 subtasks. The 26 tasks are further categorized into 4 domains, functional mobility (FM) (5 items, 28 subtasks), personal self-care (PC) (3 items, 26 subtasks), physical-instrumental activities of daily living (PIADL) (4 items, 22 subtasks), and cognitive-instrumental activities of daily living (CIADL) (14 items, 87 subtasks) (see Table 3-4). The PASS task items have standardized examiner verbal instructions and placement of task objects. The examiner rates the ADL performance of the participants based on the established criteria for each subtask. For each subtask, the independence score ranges from 0 (unable to perform task independently) to 3 (independent) (see Table 3-5), which are based on the frequency (e.g., occasional or continuous assists) and level (e.g., verbal or physical) of assistance provided by the examiner. Assistance is provided only when needed, starting from the least assistive prompt to the most assistive and intrusive prompt (see Table 3-6). The independence score for each task is obtained by averaging its subtask scores. The averaging scores that represent independence of the tasks were recoded into integer scores for Rasch analysis. In the current study, data for the PASS included only 25 tasks because the data for the oven-use task were not always available. In addition, the current study was delimited to the construct of independence, and the PASS data for safety (i.e., the personal or environmental risks when performing a task) and adequacy (i.e., the level of efficiency of task initiation, continuation, and completion, and the degree of match between the end product and criteria identified as acceptable quality when performing a task) were not included in the data analyses,.

Validity of the PASS was referenced to common geriatric ADL/IADL instruments, namely, Lawton and Brody's (1969) Scales for Instrumental Activities of Daily Living, the Older Adults Resources and Services (OARS) ADL Scale (Fillenbaum, 1988), the Comprehensive Assessment and Referral Evaluation (CARE) (Gurland et al., 1977), and the Functional Assessment Questionnaire (Pfeffer et al., 1982).. Test-retest reliability on two consecutive days was r = 0.96. Exploratory factor analysis was used to investigate the unidimensionality of the PASS independence construct. The independence scores for 26 tasks of the PASS for 1158 subjects were examined by factor analysis using SPSS 12.0, and included populations with depression, osteoarthritis, cardiopulmonary disease, and dementia, macular degeneration, stroke, and a cohort of well-elderly. The largest eigenvalue for the 26 tasks, accounted for over 37% of the variance, and was 3.44 times larger than the second largest eigenvalue, establishes unidimensionality of the PASS independence construct (Chisholm, 2005).

Table 3-4: Tasks and Domains of the Performance Assessment of the Self-Care Skills (PASS)

Functional Mobility (FM)	
	Bed transfer (move from prone to supine position and rise from bed)
	Stair use (ascend and descend stairs)
	Toilet transfer (sit and rise from a toilet)
	Bathtub/shower transfer (enter and exit tub and/or shower)
	Indoor walking (walk indoors)
Personal Self-care (PC)	
	Oral hygiene (clean teeth, dentures and/or mouth)
	Trim toenails (groom toenails)
	Dress (don and doff upper body and lower body clothing)
Physical Instrumental Activities of Daily Living (PIADL)	
	Bend, lift, and carry garbage (lift and carry garbage sack)
	Change bed linen (put on bed linens)
	Sweep (clean spillage on the floor using a broom and a dust pan)
	Clean up after meal preparation (perform clean up tasks after meal preparation)
Cognitive Instrumental Activities of Daily Living (CIADL)	
	Shop (select and purchase grocery items)
	Pay bills by check (write checks for sample utility bills)
	Balance checkbook (balance a checkbook after writing checks)
	Mail bills and checks (prepare envelopes for mailing checks)
	Telephone use (use telephone to obtain information)
	Medication management (read medication information and organize medication according to prescription)
	Obtain information: auditory (obtain information from a radio announcement)
	Obtain information: visual (obtain information from a newspaper)
	Small repairs (repair a flashlight)
	Home safety (identify and correct hazards or problems in home safety situations)
	Bingo (play bingo)
	Oven use (cook muffins in an oven)
	Stovetop use (cook soup on a stovetop)
	Use sharp utensils (cut an apple with a sharp knife)

Note. This table (Chisholm, 2005, p. 10) is used with permission.

Table 3-5: Performance Assessment of the Self-Care Skills (PASS) Independence Scoring Criteria

SCORE	CRITERIA
INDEPENDENT PERFORMANCE	
3	No assists given for task initiation, continuation, or completion
2	No Level 7-9 assists given, but occasional Level 1-6 assists given
1	No Level 9 assists given; occasional Level 7 or 8 assists given, or continuous Level 1-6 assists given
0	Level 9 assists given, or continuous Level 7 or 8 assists given; or unable to initiate, continue, or complete subtasks or task
DEPENDENT PERFORMANCE	

Note. This table (Chisholm, 2005, p. 11) is used with permission.

Table 3-6: Performance Assessment of the Self-Care Skills (PASS) Prompt Hierarchy

	LEVEL	PROMPT	DESCRIPTION
LEAST RESTRICTIVE			
VERBAL	1	Verbal support	Encouragement
	2	Verbal non-directive	Cue to alert that something is not right
	3	Verbal directive	Tell person what to do next
GESTURE	4	Gestures	Point at task object
	5	Task/environmental rearrangement	Break task down
	6	Demonstration	Assessor demonstrates/person follows
PHYSICAL	7	Physical guidance	"Hands down" – move body part into place
	8	Physical support	"Hands up" – lift body part/clothes/support
	9	Total assist	Assessor does task or subtasks for the person
MOST RESTRICTIVE			

Note. This table (Chisholm, 2005, p. 12) is used with permission.

3.3 DATA ANALYSIS

3.3.1 Data preparation

Descriptive statistics were used to describe the demographic, pathologies, and medical history related to stroke, using SPSS version 14.0. Missing data were replaced according to like subjects (i.e., LHS and RHS) and were conducted using the linear interpolation method (SPSS Inc., 2005). After replacing the missing data, Rasch analysis was performed using WINSTEPS version 3.64.1.

3.3.2 Rasch analysis

3.3.2.1 Logits of item difficulty and person ability

The ordinal data from the PASS and the BI were transformed into interval data or logits (log odds units), using Rasch analysis. The logits scale has equal unit intervals so that the difficulties of the tasks and abilities of the persons can be compared. The underlying difficulty of an item in a measurement tool is represented by item difficulty logits. Similarly, the underlying ability for performing all the tasks in a tool is represented by person ability logits.

3.3.2.2 Hierarchies of item difficulty and person ability

A hierarchy of the easiest to most difficult task items and a hierarchy of persons based on best to worst performance can be established using Rasch analysis because the items are placed on one scale and so are the persons. The current study used a score of zero as the midpoint of difficulty. For the scores from the PASS and the BI, the items with more positive logit values were harder than those with more negative values. In contrast, persons with more positive logit values had a greater abilities to perform tasks independently than those with negative logit values. Because each item of the BI had its own scaling, the Partial Credit Rasch Model (PCM) was applied, to solve the different intermediate levels that come from different numbers of responses for different items on the same instrument (Bond & Fox, 2007; Linacre, 2007).

Silverstein, Fisher, Kilgore, Harley, and Harvey (1992) pointed out that the calibration of the item difficulty logits for each item in a measure should be appropriate to discriminate person abilities. If there are huge gaps between two item difficulty logits, the measure may lose its function to discriminate various person abilities within the gaps (Silverstein, Fisher, Kilgort, Harley, & Harvey, 1992). The usefully item calibrations for judging the random variation is ±0.50 logits (Linacre, 1994).

3.3.2.3 Diagnosis: Construct validity

Fit statistics

Before interpreting the results from Rasch analysis, the data should meet the assumptions of unidimensionality (i.e., focusing on one attribute or dimension at a time) and local independence (i.e., no significant relationship between the test items, although the latent trait of the examinees is homogeneous). When the assumption of unidimensionality is met, the assumption of local independence is automatically met (Hambleton et al., 1991; Scherbaum Jr., 2003). If the data perfectly fit the Rasch model, the tool used to evaluate human performance is considered stable and reliable. If any discrepancy occurs between the Rasch model and the data collected in practice, the fit statistics provide information about how to interpret the data precisely. They are infit and outfit statistics, represented by mean squares (mean of squared residuals). For clinical observation, the reasonable task mean square error range for infit and outfit statistics is 0.5 to 1.7. Misfitting task items were with both infit and outfit mean square error values greater than the suggested level, ≤ 0.5 or ≥ 1.7 (Bond & Fox, 2007).

Principal component analysis of Rasch residuals

Principal component analysis (PCA) of Rasch residuals is an advanced method to test if the data meet the assumption of unidimensionality. It can reveal potential multi-dimensions of the measurement construct, other than the main dimension (e.g., measuring ADL functions in this study) by reporting the percentages of the variances that were explained or unexplained by the model (Bond & Fox, 2007; Linacre, 2007). Up to 5 contrasts (or additional dimensions) are identified among the unexplained variances, with eigenvalues representing the amount of variance explained by each contrast. Among the 5 contrasts, the first contrast is the largest secondary dimension of the measure and is the most important (Bond & Fox, 2007; Linacre, 2007). Rules to determine whether the whole set of data is good and represents the unidimensionality of the total measurement constructs, or whether the residuals are only considered as *noise* include (1) the explained variance of the model is greater than 60%, (2) that the unexplained variance of the first contrast is less than 5% or (3) the eigenvalue of the unexplained variance by the first contrast is smaller than 3.0 (Linacre, 2007).

Category function

Category functions (response loading for each score) of a measurement tool are identified by average measure, category fits, thresholds, and category frequency. Among these factors, average measure and category fits are the most important indicators. Average measure, the average person ability logits for the sample who respond to a particular category, is expected to increase as the category value increases. In this way, the phenomenon would fit the model's expectations that people with higher abilities are represented in the more difficult category, and vice-versa. Category fits indicate whether a category introduces noise into the measurement process. The outfit mean squares are especially important factors, which are expected to be less than 2. Thresholds are also expected to increase with the category values so that more difficult categories have larger calibration values. In addition, the range of thresholds between two categories is recommended from 1.4 to 5 logits. It indicates that the categories are distinct and lack large variance gaps between categories. Category frequency in a measurement tool is recommended to have at least 10 responses per category. A low frequency indicates that this category may be an unnecessary or redundant category.

Scale linearity

The linearity of a scale can be introduced by a plot of the raw score against the item difficulty. If the curve is closer to the ordinary least square regression line in this plot, the scale of a measurement tool is more linear (F. Wolfe et al., 2000). The R^2 linearity value is the empirical evidence of a scale's linearity.

3.3.2.4 Diagnosis: Reliability

Item reliability

Item reliability is used to determine the stability of the item difficulty hierarchy if it is applied to another sample of stroke survivors with similar traits. A higher value of item reliability also indicates that there is a wider range of item difficulty (Bond & Fox, 2007; Linacre, 2007).

Person reliability

The person reliability index is used to determine the stability of person ability ordering if these people are administered another set of items measuring the same construct. A person reliability index value is influenced by a range of ability in the sample, number of items in a test, length of the rating scale, number of categories per item, or measurement error of the test. A reliable tool is expected to have a person reliability index value higher than .80 (Bond & Fox, 2007; Linacre, 2007; Portney & Watkins, 2000). The person separation index is an alternative method to express reliability, which is used to determine stability of the person stratification levels. A person separation index higher than 2 implies good test reliability, and a 1 implies that the differences of the persons within the sample may be due to measurement error (Bond & Fox, 2007; Fisher Jr., 1992; Linacre, 2007; Wright, 1996; Wright & Masters, 1996). Furthermore, the Cronbach's alpha, analogous to the person reliability index, indicates the internal consistency of an instrument (Bond & Fox, 2007; Linacre, 2007; Portney & Watkins, 2000). A value of 0.90 implies that the instrument is considered internally consistent, and between 0.70 and 0.90 implies moderate consistency of an instrument (Portney & Watkins, 2000).

3.3.2.5 Anchoring item values

To maintain the invariance of item calibration, regardless of the intended purposes, anchoring item values should be performed before analyzing a subset of data (Bond & Fox, 2007). The item difficulty hierarchy is fixed to calculate person ability logits for a specific sample (Bond & Fox, 2007; Linacre, 2007).

3.3.3 Common person equating

To examine the hypothesis of the current study, to compare items in global (BI) and criterion-referenced, performance-based (PASS) ADL measures, the common person equating method was used. Equating in the Rasch model allows items that are administered to the same participants to link across measures, without adjusting the content. The items in the two ADL measures are then placed on a common metric, based on their item difficulty logits. That is, the hierarchy of the items in PASS and BI is presented. Thus, the two ADL measures are comparable by using this approach (Bond & Fox, 2007; Linacre, 2007; Yu & Osborn Popp, 2005).

3.4 RESULTS

3.4.1 Participants

Table 3-7 presents the characteristics of participants with stroke in the anchored database. About 60% of the participants were male and about 98% were white. The mean age was 65.35 years. There were four types of stroke in our sample, left hemispheric stroke (52.7%), right hemispheric stroke (38.2 %), cerebellum stroke (6.3%), and brainstem stroke (2.9%). The participants with ischemic stroke constituted 88.3% of the sample and those with hemorrhagic stroke constituted

11.7% of the sample. Other medical history data related to stroke status of the anchored database sample are shown in Table 3-7.

The demographics of participants in the 3 months post-stroke database are presented in Table 3-8. This sample was also predominantly male (62%), white (97.2%), and with ischemic stroke (85.9%). The mean age of this sample was 65.86 years. There was no significant difference in the number of participants with LHS (57.7%) and RHS (42.3%) ($\chi^2 = 3.38$, 1df, $p = 0.07$). Other data from medical histories or factors related to stroke are also summarized (see Table 3-7).

Table 3-7: Participants' Demographics and Stroke Characteristics of the Anchoring and 3 Month Post-Stroke Datasets

	Anchoring (N = 213)	Three Months (N = 71)
Gender		
Male (%)	58.2	62.0
Age, years (M, *SD*)	63.35 (15.24)	65.86 (13.84)
Ethnicity		
White (%)	97.7	97.2
Black (%)	1.9	1.4
Other (%)	0.5	1.4
Stroke location		
Left Hemisphere (%)	52.7[a]	57.7[c]
Right Hemisphere (%)	38.2[a]	42.3[c]
Cerebellum (%)	6.3[a]	N/A
Brainstem (%)	2.9[a]	N/A
Stroke type		
Ischemic (%)	88.3	85.9
Hemorrhagic (%)	11.7	14.1
Has prior CVA (%)	6.3[b]	8.5[d]
Has history of hypertension (%)	57.7[b]	64.8[d]
Has history of cardiac medication use (%)	35.1[b]	43.7[d]
Has history of diabetes mellitus (%)	23.1[b]	28.2[d]
Current cigarette smoker (%)	23.1[b]	22.5[d]
History of alcohol abuse (%)	6.3[b]	5.6[d]
History of atrial fibrillation (%)	13.5[b]	14.1[d]
Has used antiplates medication in week prior to stroke (%)	29.3[b]	35.2[d]

Note. [a]n = 207 for stroke location. [b]n = 208 for each characteristic. [c]No significant difference ($\chi^2_{(1)} = 3.38$, $p = .07$) between the percentages of the participants with left versus right hemispheric stroke. [d]n = 57 for each characteristic.

3.4.2 Rasch analysis

3.4.2.1 Hierarchy of Barthel Index (BI) items

To anchor the BI items for Rasch analysis, 213 administrations of the BI were analyzed. Table 3-8 summarizes the anchored item difficulty logits of the BI items for the stroke participants. Larger measure values indicate more difficult items. Bathing and the stair items were extremely difficult for a stroke survivor to perform, followed by grooming, dressing, toilet use, and feeding. The easiest items for stroke survivors were bowels, transfer (bed to chair), walking, and bladder. There were large measure gaps between the items of bathing and stairs (±0.63 logits), and between stairs and grooming (±0.59 logits). In contrast, the items of bladder and walking had the same level of difficulty.

Table 3-8: Fit Statistics and Rasch Item Difficulty Logits of the Barthel Index (BI) Items

MOST DIFFICULT				
BI items	Measure	Infit MNSQ	Outfit MNSQ	Raw
Bathing	1.80	0.67	0.57	44
Stairs	1.17	0.70	0.69	107
Grooming	0.58	1.00	0.75	60
Dressing	0.44	0.76	0.77	116
Toilet use	0.03	0.43	0.29	133
Feeding	-0.40	1.97	1.99	126
Bladder	-0.79	1.29	2.29	144
Walking	-0.79	0.72	0.68	217
Transfer (bed-chair)	-0.87	0.74	0.35	226
Bowels	-1.17	1.24	3.17	155
LEAST DIFFICULT				

Note. MNSQ = Mean Square. Raw = Sum of the scored responses to an item. Mean of the item difficulty logits = 0.00. Standard deviation of the item difficulty logits = 0.93.

3.4.2.2 BI diagnosis: Construct validity

Among the BI items, feeding and toilet use were the misfit item, in which both infit and outfit statistics were out of the recommended range (≤ 0.5 and ≥ 1.7) (see Table 3-8). The principal component analysis (PCA) of Rasch residuals for the BI revealed that the BI construct of independence was unidimensional. The Rasch model explained 82.0% of the variance and the first contrast in the residuals explained 3.5% of the variance. In addition, the factor sensitivity ratio was 0.0417, indicating that the impact of the secondary dimension was too small to influence how to measure or interpret the ADL functional status of stroke survivors. The category frequency showed that the categories of the BI functioned moderately. Generally, the average measures increased in size as the variable increased and the categories of most items with higher levels were more difficult to observe. Specifically, the item feeding had misfitting statistics (outfit statistic > 2). There were problematic categories for five items (bowels, bladder, feeding, transfer, walking), with low observed percentages. The R^2 linearity value of the BI was 0.663, indicating that the scale of the BI is somewhat linear. In summary, the construct validity of the BI for independence, based on the Rasch diagnoses, was moderate (see Chapter 2 for a more complete description of construct validity diagnosis results).

3.4.2.3 BI diagnosis: Reliability

The item reliability index of the BI was 0.92. The item difficulty hierarchy of the BI would be stable over other populations with the same traits and same sample size as our data. The person reliability index of the BI was 0.78 and the person separation index was 1.87. The Cronbach's alpha of 0.95 showed that the BI items were internally consistent. The person ability order of the BI would be stable if the sample was tested on another tool measuring the same construct of independence. The person stratification levels of the stroke participants could be moderately distinguished by the BI. In summary, the reliability of the BI for measuring independence, based on Rasch diagnoses, was moderate (see Chapter 2 for a more complete description of reliability diagnosis results)

3.4.2.4 Person ability logits of the BI

Person ability logits of the 71 participants at 3 months post-stroke for the BI were derived by using the anchored dataset. The person ability logits ranged from -4.77 to 4.74. The mean was 2.90, with a standard deviation of 2.65. The distribution curve for the person ability logits frequency of the BI was negatively skewed (see Figure 3-1) and 38 people in the sample gained the maximal person ability logits value. This indicated that there was a ceiling effect in the BI for persons at 3 months post-stroke (see Chapter 2 for a more complete description of person ability logits results)

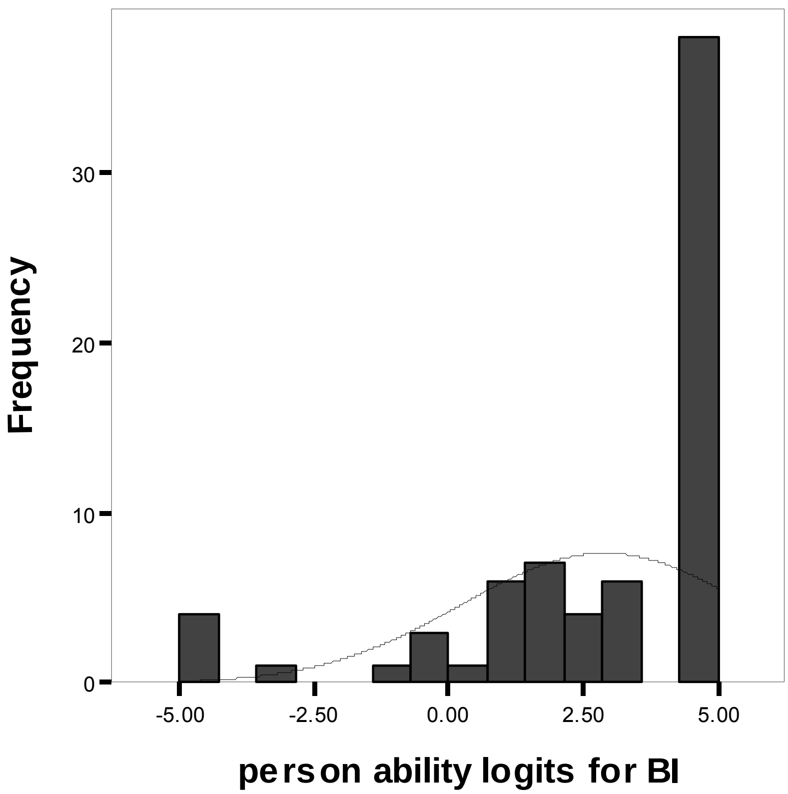

Figure 3-1: Histogram for the person ability logits frequency of the Barthel Index (BI). The curved line represents the shape of the distribution (n = 71).

3.4.2.5 Hierarchy of Performance Assessment of Self-Care Skills (PASS) tasks

To anchor the PASS task items for Rasch analysis, 211 administrations of the PASS were analyzed. Table 3-9 summarizes the anchored item difficulty logits of PASS tasks for the stroke participants. The item difficulty logits of the PASS revealed that trimming toenails and changing bed linens were the most difficult tasks for the stroke population. The financial tasks (mailing, checkbook balancing, and paying bills) were also difficult to perform, followed by a physical-emphasis IADL task, carrying garbage, and these meal preparation tasks. In contrast, bed transfer, oral hygiene, toilet use, indoor walking, and dressing were the easiest tasks to perform. In addition, bathtub or shower transfer and stair use were the most difficult tasks within the function mobility domain of the PASS. There was a large gap between the tasks of dressing and indoor walking (±0.55 logits).

Table 3-9: Fit Statistics and Rasch Item Difficulty Logits of the Performance Assessment of Self-care Skills (PASS) Tasks

MOST DIFFICULT				
PASS tasks	Logits	Infit MNSQ	Outfit MNSQ	Raw
Trimming toenails	1.27	1.48	1.29	194
Changing bed linens	1.04	1.31	1.35	229
Mailing	0.96	0.77	0.82	242
Checkbook balancing	0.92	1.00	1.04	248
Paying bills	0.68	0.89	0.90	286
Carrying garbage	0.67	1.41	1.38	288
Using sharp utensils	0.64	0.83	0.60	293
Stovetop use	0.55	0.63	0.43	307
Cleanup after meal preparation	0.49	0.65	0.43	317
Shopping	0.34	0.73	0.84	341
Sweeping	0.27	1.10	0.82	351
Medication management	0.02	0.87	1.43	387
Telephone use	0.01	0.91	0.59	388
Obtaining information-visual	-0.21	1.39	1.53	417
Playing bingo	-0.25	1.21	1.56	422
Small repairs	-0.28	0.87	0.70	426
Bathtub/Shower transfers	-0.29	1.15	1.77	427
Stair use	-0.33	1.24	1.36	432
Home safety	-0.41	0.87	0.66	440
Obtaining information-auditory	-0.52	1.28	1.36	452
Dressing	-0.62	0.77	0.73	461
Indoor walking	-1.17	1.27	0.95	501
Toilet transfers	-1.20	0.83	0.64	503
Oral hygiene	-1.22	0.66	0.47	504
Bed transfers	-1.36	0.85	0.95	511
LEAST DIFFICULT				

Note. MNSQ = Mean Square. Raw = Sum of the scored responses to an item. Mean of the item difficulty logits = 0.00. Standard deviation of the item difficulty logits = 0.74.

3.4.2.6 PASS diagnosis: Construct validity

No PASS task had both infit and outfit out-of-range values (≤ 0.5 and ≥ 1.7) (see Table 3-9). The PCA of Rasch residuals for the PASS revealed that the construct of the PASS that measured independence in stroke survivors was unidimensional. The Rasch model explained 91.7% of the variance, and the first contrast in the residuals explained 1.5% of the variance. The factor sensitivity ratio was 0.0162, indicating that the impact of the secondary dimension was too small to influence how to measure or interpret the ADL functional status of stroke survivors. According to the diagnosis of the category function, the overall categories of the PASS functioned perfectly. The scale of the PASS was linear, with an R^2 linearity value of 0.973. In summary, the construct validity of the PASS was excellent (see Chapter 2 for a more complete description of construct validity diagnosis results).

3.4.2.7 PASS diagnosis: Reliability

The item reliability index of the PASS was 0.98, indicating the item difficulty hierarchy would be stable over other populations with the same traits and same sample size as our data. The person reliability index of the PASS was 0.89 and the person separation index was 2.84, indicating the person ability order of the PASS would be stable if the sample was tested with another tool measuring independence. The person stratification levels of the stroke participants were clearly distinguished on the PASS. The Cronbach's alpha of 0.96 showed that the PASS task items were internally consistent. In summary, the reliability of the PASS was excellent (see Chapter 2 for a more complete description of reliability diagnosis results).

3.4.2.8 Person ability logits of the PASS

The person ability logits of the PASS for the 71 participants at 3 months post-stroke, based on the anchored PASS item difficulty logits, ranged from -3.79 to 4.50. The mean was 0.42, with a standard deviation of 1.73. The curve for the person ability logits frequency of the PASS showed a normal distribution (see Figure 3-2) (see Chapter 2 for a more complete description of person ability logits results).

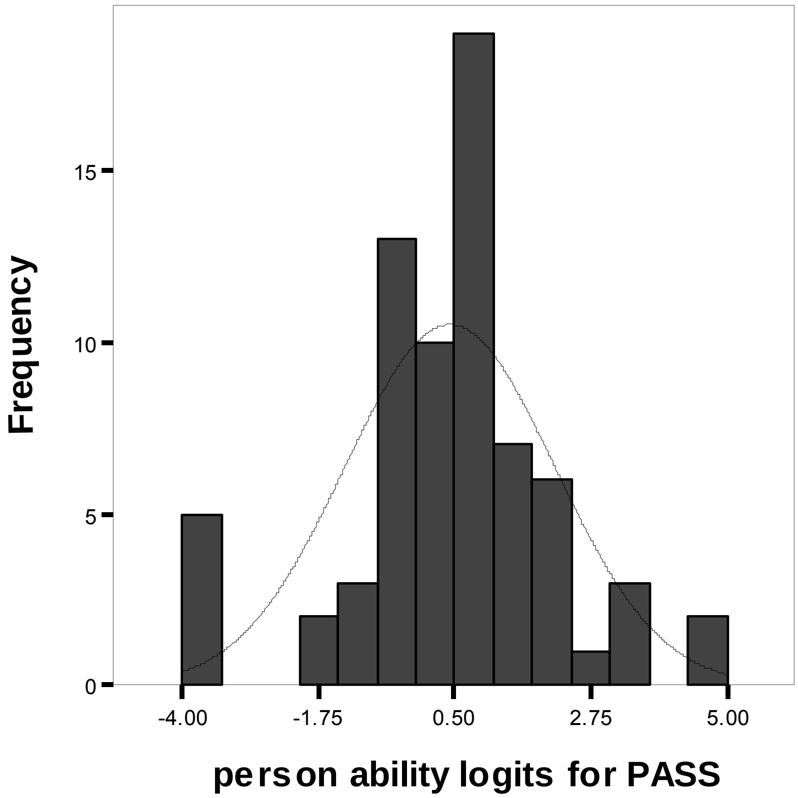

Figure 3-2: Histogram for the person ability logits frequency of the Performance Assessment of Self-Care Skills (PASS).
The curved line represents the shape of the distribution (n = 71).

3.4.3 BI-PASS equating measure

3.4.3.1 Item difficulty hierarchy of BI-PASS equating measure

Table 3-10 represents the hierarchy of item difficulty for the equating model of the BI and the PASS (the BI-PASS equating measure). For the 213 participants with stroke, the PASS task items were generally more difficult than the BI items. The trimming toenail item in the PASS was the most difficult task to perform, followed by IADL tasks, with a physical or cognitive emphasis (mailing, changing bed linens, checkbook balancing, paying bills, carrying garbage, using sharp utensils, stovetop use, cleanup after meal preparation, shopping, sweeping, telephone use, medication management, obtaining visual information, playing bingo, and small repairs). The BADL tasks in the PASS and the BI items were easier than the IADL tasks, except for the PASS task home safety and obtaining auditory information items. The easiest items were bowels, bladder, walking, feeding, and transfer from the BI. In the BI-PASS equating measure, the biggest gap appeared between obtaining auditory information task of the PASS and the bathing item of the BI (±0.39 logits), but the distance was still within the reasonable range (< 0.50 logits). Six BI and PASS items measured similar functions, namely stair use, dressing, walking, grooming, toilet use and transfers. The PASS items were all more difficult than the BI items (stair use: ± 0.62 logits, dressing: ± 0.52, walking: ± 0.98 logits, grooming: ± 0.12 logits, toilet use: ± 0.19, and transfers: ± 0.22 logits). The differences between the PASS task items and the BI items related to stair use, dressing, and walking were substaintial (> ± 0.50 logits).

Table 3-10: Hierarchy of Items in the BI-PASS equating measure

Instrument	Task/Item	Logits	Infit MNSQ	Outfit MNSQ	Raw
MOST DIFFICULT					
PASS	Trimming toenails	1.84	1.49	1.66	194
PASS	Mailing	1.67	0.93	0.82	242
PASS	Changing bed linens	1.62	1.28	1.53	229
PASS	Checkbook balancing	1.58	1.13	1.25	248
PASS	Bill paying	1.31	1.04	0.99	286
PASS	Carrying garbage	1.27	0.37	1.67	288
PASS	Using sharp utensils	1.23	0.77	0.53	293
PASS	Stovetop use	1.16	0.61	0.38	307
PASS	Cleanup after meal preparation	1.09	0.56	0.30	317
PASS	Shopping	0.90	0.93	0.80	341
PASS	Sweeping	0.88	1.02	0.73	351
PASS	Telephone use	0.62	0.92	0.58	388
PASS	Medication management	0.46	1.18	1.32	387
PASS	Obtaining information-visual	0.41	1.52	5.38	417
PASS	Playing bingo	0.35	1.31	3.66	422
PASS	Small repairs	0.28	1.00	0.80	426
PASS	Stair use	0.25	1.13	1.82	432
PASS	Home safety	0.16	0.97	0.76	440
PASS	Bathtub/Shower transfers	0.13	1.22	0.79	427
PASS	Obtaining information-auditory	0.06	1.45	2.50	452
BI	Bathing	-0.33	0.71	0.50	145
BI	Stair use	-0.37	0.83	0.56	309
PASS	Dressing	-0.50	0.96	0.73	461
PASS	Indoor walking	-0.74	1.38	1.12	501
PASS	Toilet transfers	-0.93	0.91	0.65	603
PASS	Oral hygiene	-0.97	0.81	0.55	504
BI	Dressing	-1.02	0.67	0.52	317
BI	Grooming	-1.09	0.77	0.50	161
BI	Toilet use	-1.12	0.60	0.32	335
PASS	Bed transfers	-1.44	0.95	0.82	511
BI	Transfer (Bed-chair)	-1.66	0.88	0.31	529
BI	Feeding	-1.68	1.21	1.05	328
BI	Walking	-1.72	0.74	0.47	520
BI	Bladder	-1.77	0.99	1.06	346
BI	Bowels	-1.93	1.09	1.12	357
LEAST DIFFICULT					

Note. MNSQ = Mean Square. Raw = Sum of the scored responses to an item. Analyses were performed using common person equating and partial credit model (PCM) of Rasch analysis. Mean of the item difficulty logits = 0.00. Standard deviation of the item difficulty logits = 1.14.

3.4.3.2 Diagnosis of BI-PASS equating measure: Construct validity

BI-PASS – Fit statistics

Table 3-11 also shows the fit statistics of BI-PASS equating measure. While some values of the outfit mean squares were out of the recommended range (7 PASS task items and 3 BI items), no task had both fit statistics values ≤ 0.5 and ≥ 1.7.

BI-PASS – Principal component analysis of Rasch residuals

The standard residual variance scree plot of the principal components analysis (PCA) of Rasch residuals for the BI-PASS equating measure is presented in Figure 3-3. The Rasch model explained 98.9% of the total variance (100%). If the data fit the Rasch model perfectly, the measurement dimension would explain 99.1% of the variance. The unexplained variance for the PASS-BI equating was 1.1%, and the first contrast in the residuals explained 0.2% of the variance. The eigenvalue of the first contrast was 5.4, indicating that it had the strength of about 5 items of 35 items in this analysis. The Rasch factor analysis for the first contrast yielded factor loadings, ranging from -0.58 to 0.71 (see Table 3-11). The "A" (balancing a checkbook in the PASS) and "a" (toilet use in the BI) in the plot identified the items with the most opposed loadings on the first contrast in the residuals of the BI-PASS equating measure. The tasks with positive values of factor loadings included 11 PASS task items (balancing a checkbook, bill paying, mailing, obtaining visual information, medication management, shopping, obtaining auditory information, playing bingo, home safety, telephone use, and using sharp utensils) and 1 BI item (bladder). Most of these tasks relate to complex daily tasks and require cognitive involvement, except the BI bladder item which had a factor loading value approaching zero. In contrast, the tasks with negative factor loading values included the remaining PASS tasks (toilet transfers, bed transfers, indoor walking, dressing, sweeping, stair use, oral hygiene, carrying garbage, changing bed linens, bathtub/shower transfers, cleanup after meal preparation, trimming toenails, stovetop use, and small repairs) and the other 9 BI items (toilet use, stairs, bathing, walking, transferring from bed to chair, dressing, feeding, grooming, and bowels), most of which relate to basic self-care and require physical involvement. The factor loadings against the item difficulty logits of each task in the BI-PASS equating measure were also plotted (see Figure 3-4), in which the plot represented a random pattern. The items with positive factor loadings were gathered at the more difficult level, except for one BI item, bladder (factor loading = 0.05, item difficulty logits = -1.77). Positive factor loadings also had a narrower range of item difficulty levels than those with negative factor loadings. Because there was no specific group of clustered items at a certain difficulty level in the plot, further investigations for the variance distributions were not required. In addition, the item difficulty logits yielded from the primary Rasch analysis explained 3281.3 units of the variance and the first contrast accounted for 5.4 units. The factor sensitivity ratio was 0.0016, indicating that after the Rasch measure was extracted, the first contrast influenced about 0.16% of the unexplained variance in the BI-PASS equating measure. That is, 0.16% of the unexplained variance in the BI-PASS equating measure was associated with the second largest dimension. Clinically, this impact was too small to influence how to measure or interpret independence of stroke survivors. In summary, the construct of the BI-PASS equating measure that measured ADL independence of stroke survivors was considered unidimensional.

```
      STANDARDIZED RESIDUAL VARIANCE SCREE PLOT
Table of STANDARDIZED RESIDUAL variance (in Eigenvalue units)
                                                  Empirical       Modeled
Total variance in observations     =       3316.3 100.0%          100.0%
Variance explained by measures     =       3281.3  98.9%           99.1%
Unexplained variance (total)       =         35.0   1.1% 100.0%      .9%
Unexplained variance in 1st contrast =        5.4    .2%  15.4%
Unexplained variance in 2nd contrast =        2.6    .1%   7.5%
Unexplained variance in 3rd contrast =        2.4    .1%   6.9%
Unexplained variance in 4th contrast =        2.1    .1%   5.9%
Unexplained variance in 5th contrast =        1.7    .1%   4.9%

      VARIANCE COMPONENT SCREE PLOT
       +--+--+--+--+--+--+--+--+--+
   100%+  T                       +
       |     M                    |
  V 63%+                          +
  A    |                          |
  R 40%+                          +
  I    |                          |
  A 25%+                          +
  N    |                          |
  C 16%+                          +
  E    |                          |
    10%+                          +
  L    |                          |
  O  6%+                          +
  G    |                          |
  |  4%+                          +
  S    |                          |
  C  3%+                          +
  A    |                          |
  L  2%+                          +
  E    |                          |
  D  1%+        U                 +
       |                          |
   0.5%+           1  2  3  4  5  +
       +--+--+--+--+--+--+--+--+--+
          TV MV UV U1 U2 U3 U4 U5
         VARIANCE COMPONENTS
```

Figure 3-3: Standard residual variance scree plot of principal components analysis (PCA) for the BI-PASS equating measure.

T, TV represents the total variance in the observations. M, MV represents variance in the observations explained by the Rasch measures. U, UV represents unexplained variance. 1, U1 represent first contrast (component) in the residuals. 2 (U2), 3 (U3), 4 (U4) and 5 (U5) represent second, third, fourth and fifth contrast (component) in the residuals, accordingly.

Table 3-11: Matrix of Standardized Residual Contrast 1 of PCA for the BI-PASS equating measure.

	Instrument	Tasks / Items	Loading	Measure
A	PASS	Checkbook balancing	0.71	1.58
B	PASS	Bill paying	0.65	1.31
C	PASS	Mailing	0.62	1.67
D	PASS	Obtaining information-visual	0.45	0.41
E	PASS	Medication management	0.44	1.46
F	PASS	Shopping	0.41	0.90
G	PASS	Obtaining information-auditory	0.37	0.06
H	PASS	Playing bingo	0.35	0.35
I	PASS	Home safety	0.27	0.16
J	PASS	Telephone use	0.19	0.62
K	BI	Bladder	0.05	-1.77
L	PASS	Using sharp utensils	0.01	1.23
a	BI	Toilet use	-0.58	-1.12
b	BI	Stairs	-0.58	-0.37
c	PASS	Toilet transfers	-0.54	-0.93
d	BI	Bathing	-0.53	-0.33
e	BI	Walking	-0.52	-1.72
f	BI	Transfer (bed-chair)	-0.51	-1.66
g	PASS	Bed transfer	-0.49	-1.44
h	PASS	Indoor walking	-0.48	-0.74
i	PASS	Dressing	-0.44	-0.50
j	BI	Dressing	-0.43	-1.02
k	PASS	Sweeping	-0.34	0.88
l	PASS	Stair use	-0.30	0.25
m	PASS	Oral hygiene	-0.28	-0.97
n	PASS	Carrying garbage	-0.28	1.27
o	PASS	Changing bed linens	-0.27	1.62
p	PASS	Bathtub/Shower transfers	-0.23	0.13
q	PASS	Cleanup after meal preparation	-0.13	1.09
R	PASS	Trimming toenails	-0.11	1.84
Q	PASS	Stovetop use	-0.09	1.16
P	BI	Feeding	-0.08	-1.68
O	BI	Grooming	-0.05	-1.09
N	PASS	Small repairs	-0.04	0.28
M	BI	Bowels	-0.01	-1.93

Note. Modified matrix from the output of WINSTEPS. Measure = Item difficulty logits.

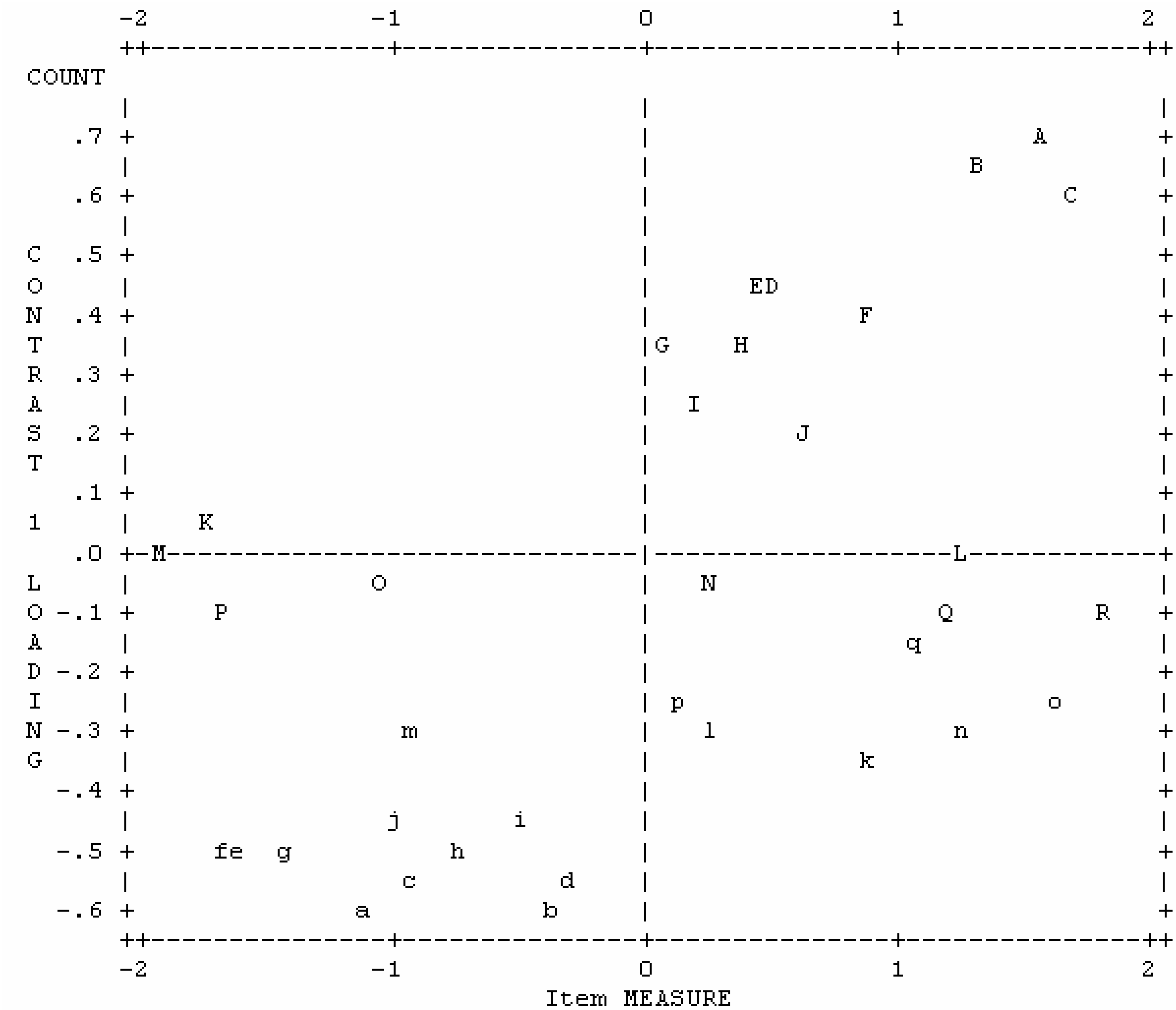

Figure 3-4: Standardized residual contrast 1 plot of PCA for the BI-PASS equating measure. The plot is plotted by the contrast 1 loadings and item measures (item difficulty logits) in the matrix of standardized residual contrast 1. This plot presents the clustered residuals visually. The “A” (feeding) and “a” (toilet use) in the plot identify the items that are the most opposed loadings on the first contrast in the residuals.

BI-PASS – Category function

Tables 3-12 and Table 3-13 present the category function of the BI-PASS equating measure after using the PCM. In the PASS, the most problematic category was the score 1 for the obtaining visual information item. The category order in the task items of bathtub or shower transfers, carrying garbage, and obtaining visual information, did not meet the expectation that more difficult items would correspond with less independent performance. Because it was a community-based sample, the non-ordered structure calibration showed that there were larger percentages of the sample scored in the extreme categories (either totally independent or totally dependent). Additionally, some categories in the stair use, bathtub or shower transfers, changing bed linens, obtaining auditory information, obtaining visual information, and playing bingo items, were out of the recommended range of the outfit statistics (i.e., outfit MNSQ > 2) (see Table 3-12). While these categories were not perfect in this partial credit model, the rating scale for the PASS as a whole still functioned well (see category function section of the PASS in Chapter 2). The BI categories generally functioned well in the current analysis, in which the observed average increased as the category value increased. Only the score 0 for the feeding item was not within the recommended outfit range (see Table 3-13).

Table 3-12: Category Functions for Items of the Barthel Index (BI) in the Partial Credit Model of BI-PASS equating measures

Item	Category	Observed %	OBSVD AVRGE	Structure calibration	Infit MNSQ	Outfit MNSQ
Bowels						
	0	5	-2.59	None	1.00	0.21
	1	3	0.23	1.02	1.19	1.59
	2	93	1.37	-1.02	1.17	0.97
Bladder						
	0	4	-2.80	None	0.60	0.25
	1	9	0.27	-0.33	1.08	1.22
	2	86	1.44	0.33	1.29	1.14
Grooming						
	0	15	-1.00	N/A	0.79	0.45
	1	85	1.54	N/A	0.71	0.75
Toilet use						
	0	7	-2.23	None	0.62	0.28
	1	9	-0.16	0.19	0.67	0.27
	2	84	1.59	-0.19	0.45	0.64
Feeding						
	0	3	-1.34	None	1.78	2.22
	1	21	0.07	-1.34	1.08	0.88
	2	76	1.55	1.34	1.13	1.07
Transfer						
	0	4	-2.68	None	1.47	1.49
	1	2	-2.03	0.21	0.24	0.04
	2	5	-0.42	0.22	0.57	0.17
	3	87	1.49	-0.43	1.03	1.06
Walking						
	0	41	-3.01	None	0.80	0.47
	1	1	-1.44	0.67	1.25	0.84
	2	13	-0.28	-1.47	0.77	0.39
	3	82	1.60	0.80	0.45	0.79
Dressing						
	0	6	-2.40	None	0.50	0.34
	1	20	0.21	-0.77	0.78	0.50
	2	74	1.71	0.77	0.64	0.77
Stairs						
	0	13	-1.04	None	1.03	0.75
	1	11	-0.14	0.42	0.65	0.33
	2	76	1.72	-0.42	0.59	0.66
Bathing						
	0	24	-0.60	N/A	0.74	0.46
	1	76	1.69	N/A	0.65	0.63

Note. Observed % = Percentage of occurrences of that category. OBSVD AVERGE = Average of the person ability logits for all persons in the sample who had that particular response to that category. Structure calibration = The difficulties estimated for choosing one response category over another. MNSQ = Mean square.

Table 3-13: Category Functions for Items of the Performance Assessment of Self-Care Skills (PASS) in the Partial Credit Model of the BI-PASS equating measure

Item	Category	Observed %	OBSVD AVRGE	Structure calibration	Infit MNSQ	Outfit MNSQ
Bed transfers						
	0	4	-2.58	None	1.36	1.13
	1	5	-1.45	-0.78	0.84	0.64
	2	11	0.56	0.37	0.95	0.84
	3	81	1.55	0.41	0.79	0.90
Stair use						
	0	20	-0.74	None	1.06	3.33
	1	5	0.62	1.33	1.49	0.82
	2	3	1.06	0.89	0.64	0.31
	3	72	1.71	-2.02	1.17	1.03
Toilet transfers						
	0	7	-2.30	None	1.01	1.12
	1	3	-0.50	0.48	1.03	0.46
	2	9	0.42	-0.33	0.80	0.56
	3	81	1.58	-0.15	0.80	0.84
Oral hygiene						
	0	7	-2.26	None	0.97	0.88
	1	2	-1.21	0.97	0.68	0.23
	2	11	0.29	-1.01	0.83	0.49
	3	81	1.60	0.04	0.65	0.81
Bathtub/Shower transfers						
	0	14	-1.31	None	0.66	0.54
	1	12	1.09	-0.20	1.73	3.59
	2	8	1.09	0.96	1.04	0.42
	3	65	1.70	-0.77	1.44	1.32
Trimming toenails						
	0	65	0.56	None	1.57	1.86
	1	1	1.60	3.31	1.56	0.13
	2	2	1.71	-0.59	1.58	0.36
	3	33	2.28	-2.72	1.41	1.97
Dressing						
	0	7	-2.17	None	0.77	0.60
	1	6	0.02	-0.29	1.08	0.71
	2	24	0.76	-0.52	0.89	0.63
	3	63	1.79	0.82	1.07	0.99
Shopping						
	0	22	-0.75	None	0.75	0.68
	1	7	0.68	0.36	0.64	0.40
	2	41	1.60	-1.55	0.65	0.65
	3	30	2.02	1.19	1.43	1.23
Bill paying						
	0	36	-0.07	None	1.02	0.89
	1	9	1.03	0.65	0.74	0.33
	2	24	1.62	-1.01	1.05	1.06
	3	31	2.22	0.36	1.12	1.39
Checkbook balancing						
	0	46	0.19	None	0.99	0.83
	1	5	1.50	1.46	1.38	0.71
	2	21	1.71	-1.63	1.08	1.07
	3	28	2.27	0.16	1.30	2.10

Table 3-13 (Continued)

Mailing						
	0	45	0.10	None	0.78	0.66
	1	5	1.23	1.32	1.18	0.58
	2	29	1.88	-2.02	0.95	0.88
	3	22	2.31	0.70	1.08	1.03
Carrying garbage						
	0	47	0.23	None	1.34	2.14
	1	1	1.83	3.91	2.63	0.23
	2	5	1.17	-2.28	1.34	1.40
	3	47	2.07	-1.63	1.38	1.42
Telephone use						
	0	28	-0.41	None	0.99	0.69
	1	2	-0.07	2.48	0.34	0.13
	2	9	1.02	-1.44	0.98	0.27
	3	62	1.91	-1.04	0.83	0.83
Medication management						
	0	16	-0.78	None	1.26	1.44
	1	5	-0.23	0.50	0.68	1.41
	2	38	1.37	-1.59	1.12	1.36
	3	41	1.87	1.09	1.25	1.12
Changing bed linens						
	0	55	0.39	None	1.26	1.60
	1	6	1.41	1.53	1.70	0.97
	2	2	1.68	0.91	0.75	0.18
	3	37	2.22	-2.44	1.28	1.89
Obtaining information-auditory						
	0	17	-0.65	None	1.55	3.30
	1	3	-0.54	1.37	0.88	1.24
	2	5	0.78	0.05	0.48	0.75
	3	75	1.65	-1.42	1.60	4.93
Obtaining information-visual						
	0	24	-0.32	None	1.44	1.87
	1	1	-1.74	2.77	0.74	9.86
	2	7	1.54	-1.55	1.63	9.32
	3	68	1.67	-1.22	1.66	6.55
Small repairs						
	0	19	-0.83	None	1.07	1.21
	1	1	-0.36	2.50	1.05	0.83
	2	17	0.92	-2.31	0.88	0.39
	3	63	1.83	-0.19	0.93	0.84
Sweeping						
	0	35	-0.16	None	1.05	0.84
	1	1	0.56	3.05	0.04	0.00
	2	7	1.33	-1.76	0.67	0.40
	3	56	0.96	-1.29	1.10	0.93
Indoor walking						
	0	8	-1.90	None	1.45	1.22
	1	5	-0.20	0.04	1.23	0.57
	2	2	1.01	1.83	0.94	2.19
	3	85	1.52	-1.87	1.53	1.12

Table 3-13 (Continued)

Home safety						
	0	18	-1.01	None	0.87	0.54
	1	2	0.11	1.80	0.40	0.07
	2	11	1.43	-1.09	1.14	1.09
	3	69	1.70	-0.71	1.18	0.99
Playing bingo						
	0	22	-0.44	None	1.34	4.68
	1	3	-0.27	1.62	0.81	2.54
	2	5	1.11	-0.08	1.26	1.09
	3	69	1.72	-1.53	1.35	5.03
Stovetop use						
	0	41	-0.09	None	0.60	0.43
	1	2	0.60	2.68	0.54	0.12
	2	12	1.53	-1.98	0.30	0.13
	3	45	2.19	-0.70	0.69	0.56
Using sharp utensils						
	0	45	0.06	None	0.79	0.50
	1	3	0.85	2.07	0.84	0.84
	2	5	1.29	-0.47	1.07	0.30
	3	47	2.19	-1.60	0.73	0.53
Cleanup after meal preparation						
	0	42	-0.08	None	0.60	0.36
	1	2	0.81	2.47	0.30	0.04
	2	4	1.29	-0.59	0.15	0.02
	3	52	2.13	-1.87	0.56	0.41

Note. Observed % = Percentage of occurrences of that category. OBSVD AVERGE = Average of the person ability logits for all persons in the sample who had that particular response to that category. Structure calibration = The difficulties estimated for choosing one response category over another. MNSQ = Mean square.

BI-PASS – Scale linearity

The scale linearity of the BI-PASS equating measure is shown in Figure 3-5. Raw data refers to the sum of the scored responses of each item (see Table 3-9). The measures were the item difficulty logits of each item. The R^2 linearity value is 0.216 and so about 22% of the variance can be explained by the model (Portney & Watkins, 2000). Therefore the scale of the BI-PASS equating measure is not linear.

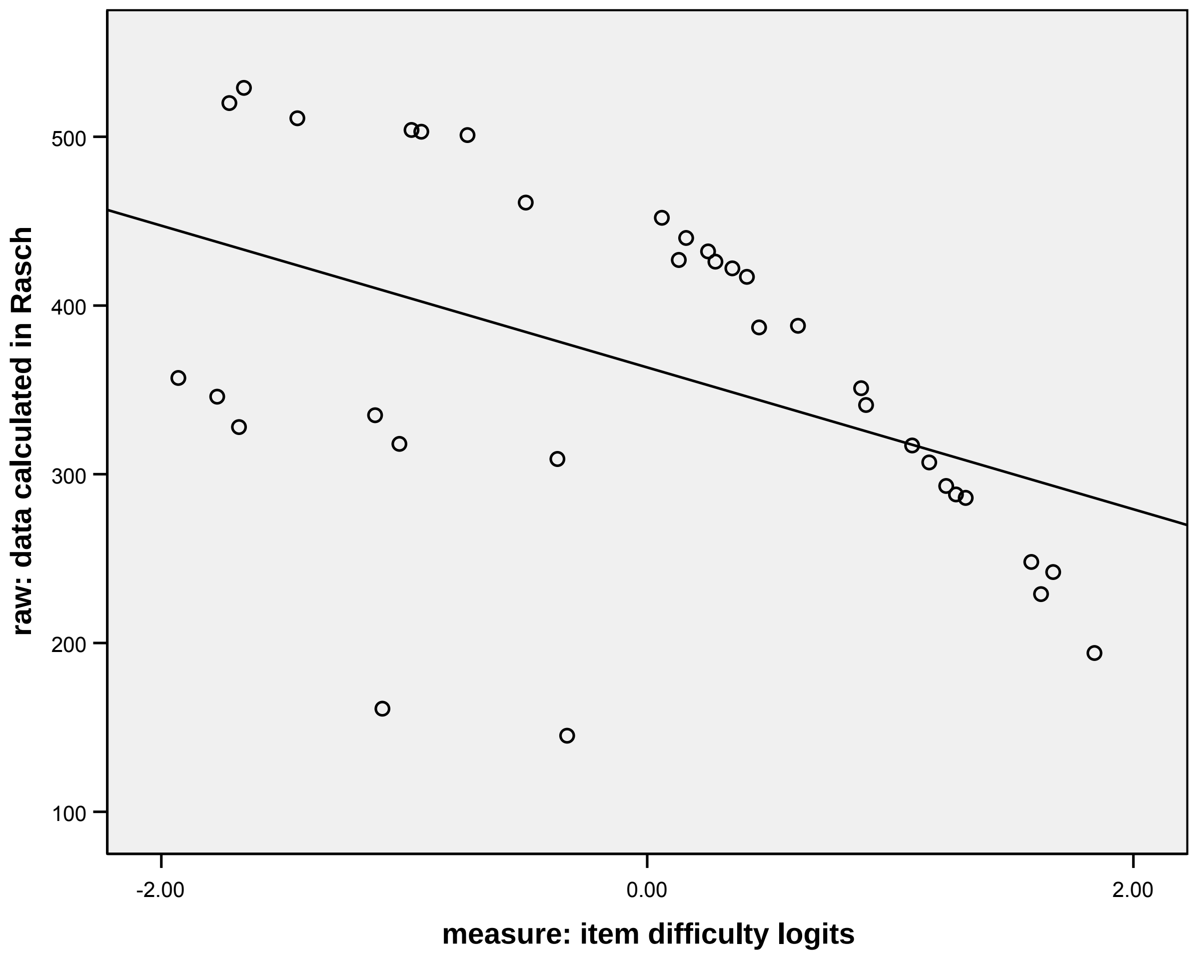

Figure 3-5: Scale linearity of the BI-PASS equating measure.
Each dot represents one item of the BI-PASS equating measure. They were plotted by their item difficulty logits (x-axis) and raw data reported in the Rasch analysis (y-axis). The line was ordinary least squares regression of y on x.

3.4.3.3 BI-PASS diagnosis: Reliability

BI-PASS – Item reliability

The item reliability index of the BI-PASS equating measure was 0.98. It implies that the item difficulty hierarchy of the BI-PASS equating measure would be consistent over other populations, with the same traits and same sample size as our data.

BI-PASS – Person separation reliability

The person reliability index of the BI-PASS equating measure was 0.92, with a good to excellent reproducibility. It implied excellent stability of person ordering. The person separation reliability of the BI-PASS equating measure was 3.42, which indicated that persons who were more independent could be reliably distinguished from those who were less independent with the BI-PASS equating measure. A Cronbach's alpha of 0.97 also showed the excellent internal consistency of the BI-PASS equating measure. Therefore, the BI-PASS equating measure is an extremely reliable tool to evaluate ADL performance for stroke survivors.

3.4.3.4 Person ability logits of the BI-PASS equating measure

After anchoring the items of the BI-PASS equating measure, the person ability logits of the 71 participants at 3 months post-stroke were calculated. The mean was 0.91 logits, with a standard deviation of 2.13. Figure 3-6 presents the item-person map with the item difficulty logit scale at the right and the person ability logit scale at the left. The distribution of the persons was generally clustered around the mean. The curve of the person ability logits distribution of the BI-PASS equating measure was normally distributed (see Figure 3-7). It implied that the items of the BI-PASS equating measure were responsive to different levels of independence for the stroke survivors at 3 months post-stroke. In addition, this map showed the interaction between persons and items. The mean of the person ability logits was around item difficulty levels of the PASS IADL task items. Also, the person who had the same person ability logit value as the item difficulty logit value of the trimming toenail task would have 50% probability of being totally independent on this task, and would also have greater than a 50% probability of being totally independent on the easier tasks.

```
     Persons MAP OF Items
       <more>|<rare>
  5      XX  +
             |
             |
             |
             |
        XXX T|
  4          +
             |
             |
             |
         X   |
             |
  3   XXXXX  +
             |
         X   |
            S|
        XX   |T
         X   |
  2     XXX  +
        XX   |  P_teonail
       XXX   |  P_linens     P_mail
      XXXX   |  P_check
   XXXXXXX   |  P_bend       P_bill
        XX   |S P_clean      P_stovetop   P_utensil
  1    XXXX M+
         X   |  P_sweep      P_shop
     XXXXX   |  P_tel
      XXXX   |  P_med
       XXX   |  P_stair      P_vision     P_repair     P_bingo
       XXX   |  P_tub        P_safety
  0     XX   +M P_aud
       XXX   |
             |  BI_Bathing   BI_Stairs
         X S |  P_dress
             |  P_walk
         X   |
 -1      X   +  P_toilet     BI_Dressing  P_oral
             |S BI_Grooming  BI_Toilet Use
             |
         X   |  P_bed
             |  BI_Feeding   BI_Transfer  BI_Walking
             |  BI_Bladder
 -2         T+  BI_Bowels
             |
             |T
        XX   |
             |
             |
 -3          +
             |
             |
             |
             |
             |
 -4          +
             |
         X   |
             |
             |
             |
 -5     XXX  +
       <less>|<frequ>
```

Figure 3-6: Item-person map for BI-PASS equating measure.
To the left of the dashed line (the logit scale) is the person ability scale, distributed from the greatest ability at the top to the least ability at the bottom. Each X represents a person in the sample. To the right of the dashed line is the item difficulty scale distributed from the most difficult item at the top to the easiest item at the bottom. Item names with “P_” in the front represent the items of the PASS. Items names with “BI_” in the front represent the items of the BI. Along the logit scale, “M” represents the mean person ability or item difficulty estimate, “S” represents the location of one standard deviation (SD) from the mean estimates, and “T” is the second SD away from the mean estimates.

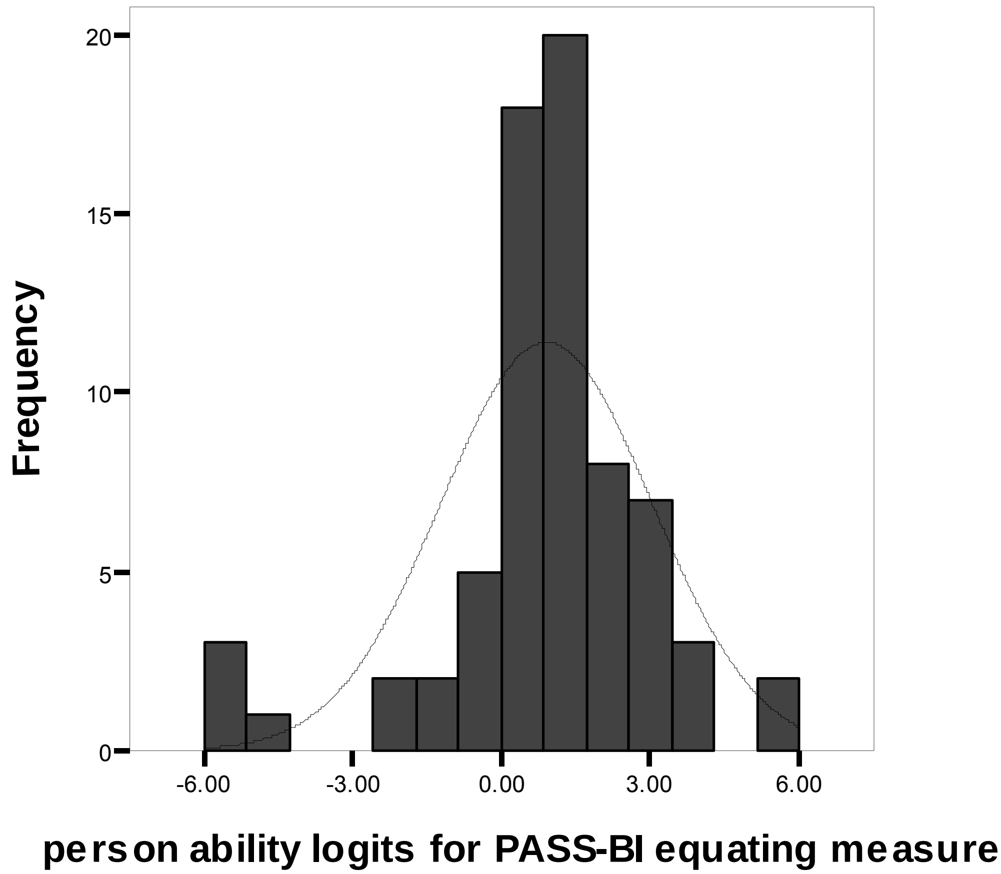

Figure 3-7: Histogram of the person ability logits frequency for the BI-PASS equating measure. The curved line represents the shape of the distribution (N = 71).

3.4.4 Common person equating

The common person equating was performed to compare items in the BI (a global ADL measure) and the PASS (a criterion-referenced ADL measure). Figure 3-8 is a scatterplot which also illustrates the comparison of both measures. The person ability logits for each participant who performed both PASS and BI were plotted. Each dot represents a person's person ability logits for the PASS and BI. Generally, a person who performed more independently on the BI also performed more independently on the PASS. Only 3 people did not fit this pattern, and the dots representing their person ability logits for the PASS and BI were not within the confidence intervals. An identical equivalence line in the plot represents the error-free Rasch modeled equivalence between the person estimates. The equating logits, an interval from the intersection of the empirical equivalence line and the BI (x-axis) to the intersection of the BI (x-axis) and the PASS (y-axis), is about 2.5 for the two measures, which indicates that the PASS is about 2.5 logits more difficult than the BI for stroke survivors at 3 months post stroke (Bond & Fox, 2007; Linacre, 2007).

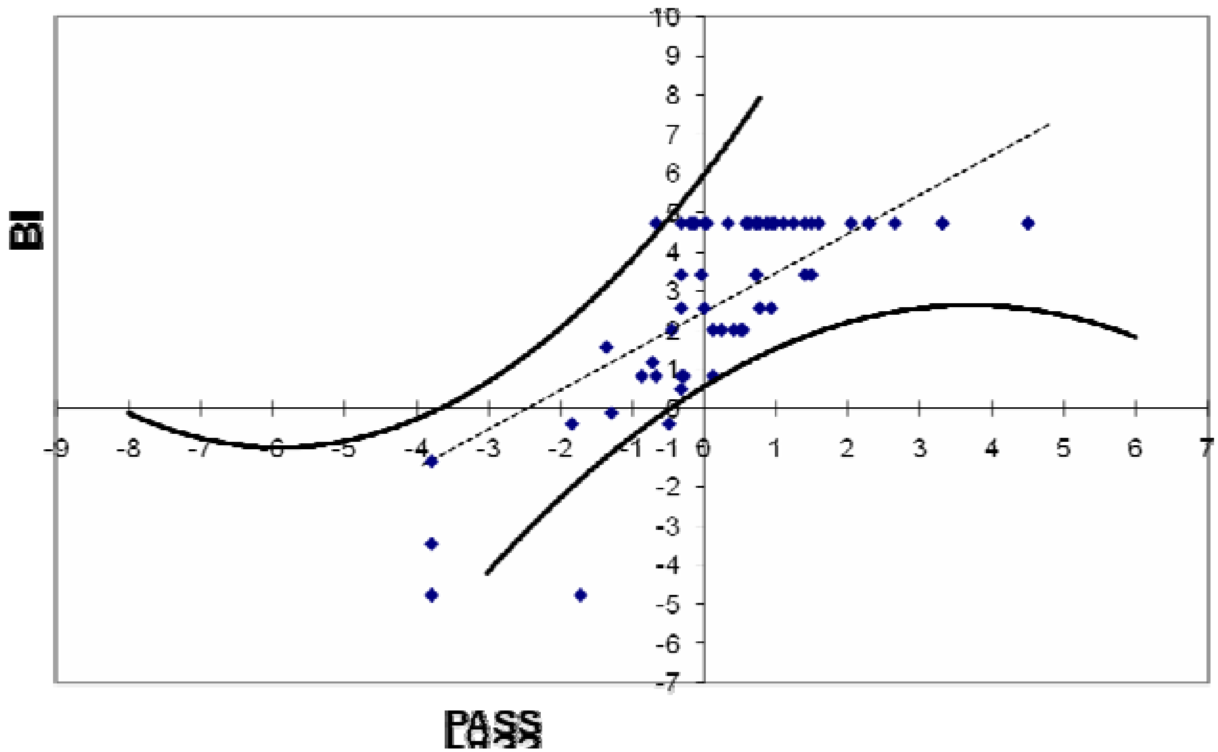

Figure 3-8: Scatterplot of person ability logits of Performance Assessment of Self-care Skills (PASS) and Barthel Index (BI). The dashed line represents an identical equivalence line. The two curved lines represent upper and lower two-sided 95% confidence intervals. Each dot represents each participant, plotted according to their person ability logits of PASS against person ability logits of BI.

3.5 DISCUSSION

The current study compared the BI and the PASS functional assessment measures for stroke survivors at 3 months post-stroke. Even though the BI and PASS were discussed in the previous chapter, in the current study the global non-summative measures (the GOS5, GOS8, and mRS) were extracted. We then compared the items of the two rehabilitation-relevant measures (the BI and PASS) on a single metric using the Rasch analysis partial credit model (PCM) and common person equating. Although both tools measure the construct independence, our findings supported the hypothesis that, overall, the PASS was more difficult than the BI for participants at 3 months post-stroke.

The construct validity and reliability of the PASS and the BI were confirmed separately in a previous study (see Chapter 2). In the current study, the BI-PASS equating measure was diagnosed as a unidimensional, valid, and reliable tool. However, the scale was no longer linear, which may reflect combining the moderately linear BI scale, with the strongly linear PASS scale. Another potential explanation for lack of linearity in the BI-PASS equating measure is the method of obtaining the independence data for current study. BI data were obtained by self-report, while the PASS data were gathered by performance-based observations. It is not unusual for subjects to self-report better performance than what is observed during performance testing. Moreover, in addition to BADL items, the PASS includes 13 IADL items, and the 10 BI items only focus on BADL. Furthermore, if the BI is used by practitioners as a screening tool in the clinic to decide whether, and what, further evaluations and interventions should be addressed, the Rasch analysis item-map data indicate that practitioners would incorrectly assume that the patients were independent, and might not refer them for further evaluation or intervention. Self-reporting level of independence on the BI items not only indicated that stroke survivors overestimated their abilities, but it may also have revealed that they were unaware of their disabilities.

Based on item hierarchy of the BI-PASS equating measure, most BI items were ranked at a less difficult level than the PASS tasks. The 10 BI items consist of BADL, which are generally considered to be easier to perform (Lawton, 1983). The 25 PASS items include 8 items which are considered BADL and 13 items which are IADL, in which the BADL items were generally easier than the IADL items on this metric. In addition to the item hierarchy, the item-person map showed that at 3 months post-stroke, independence was "easier" to achieve for BI items than for PASS items. Similarly, when we examined person ability on the item-person map, it showed that stroke survivors who were extremely independent on the BI were only moderately independent on the PASS. The common person equating measure also indicated that overall, the PASS is about 2.5 logits more difficult than the BI, and the chance for people at 3 months post-stroke to be totally independent on PASS items was much less than on BI items. For stroke survivors who were discharged and lived in the community as did our sample, to be independent when performing IADL tasks would be more challenging than when performing BADL tasks, such as the BI items and the motor items of the FIM. Our findings suggested that assessing IADL tasks is essential for stroke survivors, especially when they live in community.

One of the most clinically relevant findings of the current study was that when the BI and PASS items measured similar functions, the PASS items were always more difficult (i.e., stair use, dressing, walking, transfers, toilet use, grooming). Among these items (e.g., dressing, stair use, and walking), most of their differences in item difficulty logits were substaintial (i.e., differences between the two item difficulty logit values of over 0.50 logits). These differences indicate that there were clinically relevant differences between the two tools in how they measure independence, and for some items the difference is substaintial. Because of self-reported overestimation of abilities, practitioners should be careful when interpreting self-report data from the BI. For example, on the item-person map, a patient who reports being independent on the BI stairs item, would have a greater chance of being dependent on all items above the BI stairs item on the item difficulty hierarchy, which would include the PASS stair use task as well as all the IADL tasks. However, if the BI self-report were used as a screening tool, and if the same patient wanted to be independent in stairs, bathing, and financial management, the item difficulty hierarchy guides the practitioner to further evaluate stair use with the performance-based PASS, as well as evaluate tub transfers and several other items located on the hierarchy between the BI stair item and the PASS financial management items.

The item-person map of the BI-PASS equating measure showed that majority of the participants at 3 months post-stroke were challenged by the IADL task items in the PASS. This was especially true for the financial tasks (mailing, checkbook balancing, and paying bills) and the meal preparation tasks (using sharp utensils, stovetop use, and cleanup after meal preparation), although these tasks did not require a great amount of physical demands and were clustered as cognitive-oriented tasks in the PCA. Cognitive dysfunctions, such as memory, abstract reasoning, and problem-solving impairments, generally are associated with the performance of BADL (Carter, Oliveira, Duponte, & Lynch, 1998), but do not always statistically influence BADL (Zinn et al., 2004). In addition, stroke survivors with cognitive impairments had worse recovery of IADL tasks than did those without cognitive impairments (Zinn et al., 2004). The hierarchy developed by Lawton (1983) demonstrated that financial tasks were harder than general BADL and IADL tasks (Dittmar, 1997). The complexities of the financial tasks involved a greater amount of memory, abstract reasoning, and problem-solving skills. Hoskin, Jackson, and Crowe (2005) also reported that patients with acquired brain injury showed dysfunctions in memory and executive functions had the poor abilities on novel tasks, such as money management. Furthermore, the meal preparation required managing multiple implements in the kitchen in an adequate and safe manner, and stroke survivors required more assistance for these tasks than for the BADL tasks, which was also noted in a study by Hartman-Maeir, Soroker, Ring, Avni, and Katz (2007). In addition to the challenge of IADL tasks, trimming toenails was the most difficult task for stroke survivors at 3 months post-stroke. This task is usually considered a BADL, and less difficult than IADL tasks. However, this task was not included in any other ADL measure, nor was there any reference to trimming toenails in the literature. This information is thus new to the body of knowledge. Stroke survivors usually present with a certain degree of physical impairment, such as hemiplegia or synergy patterns, as a consequence of the stroke. These impairments can influence their abilities to complete the activities without any assistance (Ashburn, 1997; Woodson, 2002). To be independent in the trimming toenails task, the person was required to flex the upper trunk, which could evolve the synergy patterns and then limit their abilities to reach their feet and maintain balance at the same time.

Limitations and Recommendations

The current study had some limitations. Because this was a secondary data analysis, our sample was not large enough to distinguish between participants with ischemic and hemorrhagic strokes or with specific brain lesion locations. This may have yielded greater variances because the mechanisms of the strokes usually influence the severity of the disabilities in the stroke population . In addition, the BI data were collected by self-report and the PASS data were collected by performance-based observation. These different strategies might influence the interpretation of the results. However, the intention of the current study was not only to compare the two instruments, but also to utilize the BI-PASS equating measurement. Future studies are recommended that include a more homogeneous sample at 3 months or more post-stroke, and one which focuses on the PASS subtasks, so that the item hierarchy can be even more specific.

3.6 CONCLUSIONS

The current study compared item difficulties from the rehabilitation-relevant PASS and BI measures with stroke survivors at 3 months post-stroke. The PASS psychometrics indicated that the PASS was more unidimensional, valid and reliable than the BI for evaluating the ADL functional status of stroke survivors. Rasch analysis provided evidence that the PASS was not only more difficult than the BI, but also more able to discriminate between the functional abilities of stroke survivors more discretely. The hierarchy of BADL and IADL tasks of the BI-PASS equating measure visually illustrates how difficult IADL items are compared to BADL items, and provides evidence to guide more in-depth evaluation and intervention priorities, based on item difficulty. The item-person map of the BI-PASS equating measure also shows how the person ability and item difficulty interact and how important it is to assess independence on the IADL items for stroke survivors at 3 months post-stroke. This information will be useful for clinical practitioners as well as stroke survivors, so that interventions can be planned appropriately.

4.0 COMPARING THE IMPACT OF LEFT AND RIGHT HEMISPHERIC STROKE ON ACTIVITIES OF DAILY LIVING

4.1 BACKGOUND

4.1.1 Impairments related to stroke locations

Stroke is a major cause of disability in the United States (American Heart Association, 2007). The lesions that cause loss or abnormality in the brain determine the behavioral sequelae and prognoses of stroke survivors. In general, the common motor and cognitive impairments associated with stroke in the left hemisphere (LHS) are apraxia, motor sequencing deficits, and aphasia; whereas with stroke in the right hemisphere (RHS) they are neglect phenomena, visual spatial syndromes, and postural instability (Forerch et al., 2005; Harrington & Haaland, 1992; Mills & DiGenio, 1983; Stone, Halligan, & Greewood, 1993). These impairments of LHS and RHS have been widely studied, and their contributions to disabilities have been demonstrated (Ferrucci et al., 1993; Lind, 1982). However, the degree of impairment does not always predict the level of disability (Nagi, 1965; Roth et al., 1998). Clinical observation and scientific investigation have found that reduced impairments do not fully explain reduced disabilities in stroke survivors during rehabilitation, even when they are significantly correlated (Brandstater, 1990; Ferrucci et al., 1993; Roth et al., 1998). Whereas previously the emphasis was on assessing impairments of stroke survivors, more recently, the emphasis has changed to measuring disabilities (Lindeboom et al., 2003).

Therefore, this chapter begins with a systematic review of the literature, which describes differences in disabilities, between persons with LHS and RHS. The systematic review is followed by a secondary data analysis of 71 stroke survivors. Using Rasch analysis, we aimed to compare the functional impact of LHS and RHS on disability, at 3 months post-stroke. It was intended that the findings from the study would provide practical guidance related to intervention planning for task disabilities associated with LHS and RHS.

4.1.2 Disabilities of stroke survivors

The disabilities related to stroke include limitations in the "capacity to carry out any activity within the normal range achievable by a human being," (Carod-Artal, Gonalez-Gutierrez, Herrero, Horan, & de Seijas, 2002, p. 207) such as eating, dressing, and financial management. Rehabilitation practitioners are concerned with how these patients can perform everyday tasks despite their underlying impairments (Bernspang & Fisher, 1995). Stroke survivors expect that following rehabilitation, they will not only have a return of their functional abilities, they will be able to return to their former activities (Hafsteinsdottir & Grypdonck, 1997). To identify a patient's ability to perform daily activities, clinical practitioners rely on assessments that measure the functional status (Rogers, 1983). These assessments allow clinical practitioners to observe patients performing everyday activities to "help to explain, confirm, or cast doubt" (Rubenstein, Calkins, Greenfield, et al., 1988, p. 563) on their functional status.

4.1.3 Stroke survivors and functional status

Functional status refers to a person's ability to perform activities of daily living (ADL) and fulfill social roles, at a specific point in time (Rubenstein et al., 1988). Duncan, Jorgensen, and Wade (2000) suggested that ADL should be the primary functional status measure in stroke rehabilitation due to their relative objectivity, simplicity, and relevance to patients. However, determining what parameter of functional status is measured can be somewhat ambiguous. One common approach used in rehabilitation is to evaluate independence in ADL (Pamela W. Duncan et al., 2000; Rogers & Holm, 1998; van Boxel et al., 1995). These functional status data, represented by independence, provide a means for rehabilitation practitioners to communicate, document progress, and compare or monitor the functional status of stroke survivors (Portney & Watkins, 2000; Rubenstein et al., 1988). Functional status data significantly influence the rehabilitation practitioner's understanding of a stroke survivor's disability and are often a key element in their clinical reasoning as they develop treatment plans (Dittmar, 1997; Rogers, 1983). The data provide the most concrete and detailed information about the disabilities of stroke survivors (van Boxel et al., 1995). In rehabilitation, it is extremely important to have an understandable measure that both describes the functional status of a patient and guides intervention (Smith & Taylor, 2004).

4.1.4 Issues related to ADL functional status measures

4.1.4.1 ADL classification

There are two fundamental domains for classifying ADL tasks (Reuben & Solomon, 1989). The first domain is basic ADL (BADL), which refers to the self-maintenance activities involving functional mobility and personal care. Ambulation or wheelchair mobility, transfers, feeding, hygiene, toileting, and bathing are examples of BADL tasks (Rogers & Holm, 1998). The second domain is instrumental ADL (IADL), which are more complex activities that are home management tasks and tasks required for independent living in the community. Shopping, cooking, housekeeping, laundry, use of transportation, managing money, managing medicine, and use of the telephone are examples of IADL tasks (Lawton & Brody, 1969). In addition, BADL can be further delineated into functional mobility and personal care and IADL into those with a cognitive (CIADL) or physical emphasis (PIADL) (Holm & Rogers, 1999).

4.1.4.2 Assessment method

There are two major assessment methods to evaluate ADL functions, informant-based and performance-based observational (PBO) formats (Rogers & Holm, 1998). Informant-based assessments apply the method of asking questions via interviews or questionnaires, such as the Barthel Index (BI). PBO assessments use the approach of observing clients' ADL performance in natural or laboratory conditions, such as the Performance Assessment of Self-Care Skills (PASS). PBO assessments tend to be more valid and reliable than informant-based assessments (Rogers & Holm, 1998; Sinoff & Ore, 1997). Moreover, PBO assessments allow researchers or rehabilitation practitioners to gather more detailed and specific information about the processes of executing ADL tasks (Finlayson et al., 2003).

4.1.4.3 Scoring systems

ADL functional status scoring systems can be global or task specific. For a global non-summative measure, an ordinal single scale is used. This scale reflects increasing or decreasing levels of overall functional status. The modified Rankin Scale (mRS) and the Glasgow Outcome

Scale (GOS) are examples of global non-summative measures. The second type of global scoring system uses summative scoring where the total score is obtained by adding up the obtained ratings on all items. The Functional Independence Measure (FIM) uses global summative scoring (Ravaud et al., 1999; Ring et al., 1997; van Hartingsveld et al., 2006). The summative scores indicate the severity of disability, however, rehabilitation practitioners can only obtain a broad impression of the patient's performance (e.g., a higher score on the BI indicates "more independence in BADL"), and the exact functional ability of patients cannot be determined, thus decreasing these measures' usefulness. Global averages , the third type of global scoring system, uses the grand means of like items, indicating severity of disability for specific ADL domains (Holm & Rogers, 1999; Rogers et al., 2003). This method provides greater direction to rehabilitation practitioners, because the ADL domains are derived from like tasks (e.g., CIADL). The Performance Assessment of Self-Care Skills (PASS) is an example a functional status measure that uses ADL domain averages (Holm & Rogers, 1999).

Task specific scores are derived from the ratings of single tasks with critical and essential standards (i.e., criterion-referenced tests). The PASS also uses a task specific scoring system, in which each task contains subtasks that are sequenced and performance criteria that are rated. The data from a task specific scoring system has clear criteria for rating performance, each item can stand alone, and inter-reliability is available for all items as well as the total tool (Rogers et al., 2001). Task specific measures also provide meaningful functional status data for rehabilitation practitioners because they identify to what degree the patient can perform specific daily task independently. In addition, although the summative scores or task average scores for a whole tool or ADL domain were useful for policy formation, clinical practitioners provide task-specific interventions. Therefore, a valid and reliable task-specific score is essential for clinical practice (Raina, 2005).

4.1.5 Systematic review: Evidence on ADL performance of LHS and RHS survivors

To better understand the differences in ADL functional status between patients with LHS and RHS, we conducted a systematic review of the literature. The reviewed literature addressed the differences in ADL functional status between patients with LHS and RHS.

4.1.5.1 Methods of systematic review

Search strategies

Research studies were identified initially through electronic database searches of MEDLINE (1966 to present), Cumulative Index of Nursing and Allied Health Literature (CINAHL) (1982 to present), PsychINFO (1967 to present), and the Cochrane Database. Keywords were left hemisphere cerebrovascular accident (stroke) and right hemisphere cerebrovascular accident (stroke), combined with functional status, activities of daily living, or disability outcomes. Further searches were conducted for authors who have pursued comparison of LHS and RHS, and hand searches were conducted of journals that have included such studies. Review of citations in retrieved articles and consultation with known experts in the field of stroke functional status study were also used to search for additional articles.

Article criteria

Criteria for the literature search were set in advance. Articles had to meet the following criteria: (a) Sample diagnosed with stroke or cerebrobascular accident (CVA); (b) ADL assessments were included; (c) comparisons of functional status between patients with LHS and RHS were included; (d) definitions of target outcomes were identified; (e) articles were in English.

Level of evidence

The literature was ranked by the level of evidence. The hierarchy of levels of evidence was based on the classification system created by Moore, McQuay, and Gray (1995). This designation of rank helps clinical practitioners to select the best evidence that guides clinical practice for a certain population or patient. Based on the study design, there are five levels of evidence: (a) Level I is "strong evidence from at least one systematic review of multiple well-designed randomized trials." (b) Level II is "strong evidence from at least one properly designed randomized controlled trial of appropriate size." (c) Level III is "evidence from well-designed trials without randomization, single group pre-post, cohort, time series, or matched case-controlled studies." (d) Level IV is "evidence from well-designed nonexperimental studies from more than one center or research group." (e) Level V is "opinions of respected authorities, based on clinical evidence, descriptive studies, or reports of expert committees" (Holm, 2000, p. 576). The reviewed literature was classified by these criteria to help readers identify the best evidence.

4.1.5.2 Results of the literature search

Twelve studies, which compared the impact of LHS and RHS on ADL functional status, were found. Seven articles were categorized as Level III evidence, three articles were Level IV evidence, and two articles were Level V evidence. Table 4-1 includes summaries of the content of each article. Because the primary purpose of each article was not the same, the main results of each article did not always report comparisons of ADL functional status for stroke survivors with

LHS and RHS. Thus, findings related to the theme of the current review were emphasized in italics.

Among the 12 reviewed articles, some of the measurement tools did not measure task disabilities (i.e., the Assessment of Motor and Process Skills (AMPS), the Jebsen-Taylor Hand Function Test (JHFT), and the Neurobehavioral Specific Impairment (NSIS) subscale of the Arnadottir-OT-ADL Neurobehavioral Evaluation (A-ONE)). In addition, the cognitive items of the FIM measure impairment rather than task disability. Data from these assessments could not be interpreted by "task" per se. Therefore, the following review only delineates those findings yielded from the assessments that measured task disabilities, namely, the FIM-total, the FIM-motor, the FIM-gain (changes in the FIM-total scores over time), the Functional Independence Scale (FIS) of the A-ONE, the Klein-Bell Activities of daily living, and the two non-standardized clinical assessments. The FIM evaluated motor tasks in self-care, sphincter control, and transfer locomotion. The FIS of the A-ONE included BADL tasks of dressing, grooming and hygiene, transfer, and feeding domains, but not the communication domain (Gardarsdottir & Kaplan, 2002). The Klein-Bell Activities of daily living included 5 BADL (dressing, elimination, mobility, feeding, and bathing/hygiene) and 1 IADL (i.e., the emergency telephone communication) task (Shiotsuka, Burton, Pedretti, & Llorens, 1992). The ADL Health Index in the Johansson, Jadback, Norrving and Widner article (1992) measured both BADL and IADL tasks (i.e., dressing getting in and out of bed, cutting hard food, walking, washing, lavatory visits, picking up things from the floor, and going out along). The ADL functional status measures in the Mills and DiGenio article (1983) evaluated one BADL task (dressing). .

4.1.5.3 Differences in ADL functional status between patients with LHS and RHS

Dependence in daily living tasks among stroke survivors has been reported in the literature. Most comparisons of global ADL performance on the FIM-total and on the total scores of the Klein-Bell Activities of Daily Living Scale between patients with LHS and RHS found similar performance or lack of significant differences between the two groups at admission and discharge for each study, as well as for the FIM-gain (Ring et al., 1997; Yavuzer, Kucukdeveci, Arasil, & Elhan, 2001). However, Yavuzer et al. (2001) found that patients with LHS were significantly more independent on the FIM-total at admission that those with RHS, but the two the groups were similar at discharge. Chae and Zorowitz (1998) found the opposite --- that patients with RHS were significantly more independent on the FIM-total at admission, but patients with LHS improved significantly more than those with RHS, and thus the two the groups were similar at discharge. Also, there was a significant difference in the shave/make-up task, under the grooming domain of the A-ONE FIS, and patients with RHS were more independent than those with LHS (Gardarsdottir & Kaplan, 2002).

Although still not significantly different, findings generally indicated that the performance of patients with LHS was generally better those with RHS (Chae & Richard, 1998; Gardarsdottir & Kaplan, 2002; Johansson, Jadback, Norrving, & Widner, 1992; Ring et al., 1997; Shiotsuka et al., 1992; Yavuzer et al., 2001). Chae and Zorowitz (1998), who examined performance based on cortical versus subcortical lesions found that patients with LHS and cortical lesions were slightly more independent than those with RHS and cortical lesions on mobility and locomotion domains at admission and on self-care, mobility, and locomotion domains at discharge. Moreover, patients with LHS and subcortical lesions were slightly more independent than those with RHS and subcortical lesions on self-care and sphincter domains at admission and on sphincter and locomotion domains at discharge. Johnansson, Jadback, Norrving, and Widner (1992) reported that patients with LHS were consistently more independent than RHS over time (pre-admission, and 48-hours, 6-months, 12-months after admission) on the items in their non-standardized clinical ADL measure. Moreover, Heinemann et al. (1993) and Tsuji, et al. (1995) discussed FIM item difficulties for patients with LHS and RHS using Rasch analysis. Only results from the Heinemann et al. study were further analyzed, because the methods to obtain item difficulty logits in the Tsuiji et al. study are not valid using current methods (i.e. the item difficulty was not anchored). In contrast to other studies, Heinemann et al. found that FIM feeding, grooming, bowel, and bladder items were more difficult for patients with LHS than those with RHS. Statistical differences of item difficulty between LHS and RHS were not reported. However, the FIM items were anchored on the performance of multiple diagnostic populations not just stroke survivors. In contrast, with the six Klein-Bell Activities of Daily Living Scale subtasks, patients with RHS were slightly more independent than those with LHS at admission and discharge (Shiotsuka et al., 1992). Similarly for most tasks on the FIS of the A-ONE (don/doff socks, comb hair, maneuver, drink, finger feeding, use fork or spoon, and use knife), patients with RHS were slightly more independent than those with LHS (Gardarsdottir & Kaplan, 2002). However, some studies found that the two groups were not statistically different in their performance of some items, namely putting on/taking off pants (FIS of the A-ONE; Gardarsdottir & Kaplan, 2002) and the bathing, locomotion, and tub transfer items of the FIM (Heinemann et al., 1993).

In summary, most results revealed that the ADL performance of patients with LHS and RHS was not significantly different at admission or discharge. When significant differences were reported (Chae & Richard, 1998; Gardarsdottir & Kaplan, 2002; Yavuzer et al., 2001), patients with LHS were more independent than those with RHS on global functional status measures (i.e., the FIM-total and the FIM-gain) at the overall level. No study was found which compared differences between patients with LHS and RHS at the IADL domain or task levels that included IADL. Therefore, the current study compared LHS and RHS survivors, at 3 months post-stroke, on a valid and reliable performance-based observational tool (the Performance Assessment of Self-Care Skills (Rogers & Holm, 1984), at three levels: overall, domain (functional mobility, personal self-care, IADL with a cognitive emphasis, IADL with a physical emphasis), and task.

Table 4-1: Evidence of ADL Performance on LHS and RHS survivors

Authors	**Purpose of the study**	**Inclusion and exclusion criteria**	**Level of evidence**	**Independent variable**	**Dependent variable**	**Characteristics of subjects**	**Data analysis** ------------ **Relevant findings**
Rexroth, P., Fisher, A. G., Merritt, B. K. & Gliner, J. (2005)	To determine if people with a CVA differ in their abilities to perform ADL tasks, by side of lesion, gender, or age	Inclusion: - Existing data from AMPS database - LHS and RHS - Performed 2 AMPS tasks - 45 years old or older - had previously or were currently receiving occupational therapy services Exclusion: - scored by more than 10 raters - rater scoring error: high AMPS motor > 4.0 logits; process > 3.0 logits	**III**	Side of lesion - LHS - RHS Gender - Male - Female Age group - 45-59 - 60-64 - 65-69 - 70-74 - 75-79 - 80-84	AMPS - Motor skill - Process skill	LHS: - n = 1939 - Male: n= 954 - Female: n = 985 RHS: - n = 1970 - Male: n= 954 - Female: n = 985 Whole sample: - Age 45 – 94 yr - Ethnicity: white, black, Hispanic, Asian, and other	**ANOVA** AMPS motor skill: - *LHS significantly more impaired than RHS in motor ability (p = .02, d = - .09)* - Men significantly less impaired than women in motor ability (p = .02, d = - .17) - Motor ability declined as age increased - Significant differences (greater impairment) in motor ability among age groups that differed (increased) by 10 years or more (p = .01), except for adjacent age groups of 45 to 59 years and 60 to 64 years (d = .19) AMPS process skill: - *LHS significantly more impaired than RHS in process ability (p = .01, d = .10)* - Women significantly less impaired than men in process ability (p = .01, d = .07) - Process ability declined at a faster rate as age increased - Significant differences (greater impairment) in process ability among age groups that differed (increased) by 10 years or more (p = .01) **Many-faceted Rasch (MFR)** AMPS motor skill: - No clinically detectable difference (± 0.43 logits) in skill item calibrations between LHS and RHS: range from -0.14 to 0.13 logits AMPS process skill: - No clinically detectable difference (± 0.43 logit) in skill item calibrations between LHS and RHS: range from 0.17 to 0.29 logits

Note. CVA = Cerebrovascular accident. ADL = Activities of daily living. AMPS = Assessment of motor and process skills. LHS = Left hemispheric stroke. RHS = Right hemispheric stroke.

Table 4-1 (Continued).

Authors	Purpose of the study	Inclusion and exclusion criteria	Level of evidence	Independent variable	Dependent variable	Characteristics of subjects	Data analysis ---------- Relevant findings
Wetter, S., Poole, J. L., & Haaland, K. Y. (2005)	To investigate the functional impact of ipsilesional motor deficits after unilateral stroke and the best predictors of those deficits	<u>Inclusion:</u> - Right handed - Able-bodied participants tested with left (LAB) and with right hand (RAB) - LHS and RHS <u>Exclusion:</u> - JHFT performance more than 2 standard deviations below the mean of the respective group	**III**	<u>Side of lesion</u> - LHS - RHS <u>Brain lesion</u> - Stroke - Able-bodied <u>Apraxia</u> - With apraxia - without apraxia	<Functional skills> <u>JHFT</u> (Longer times represent greater upper extremity impairments) - Total time - Writing time: to detect the impact of aphasia <Neuro-psychologic variables> <u>Ideomotor limb apraxia</u> <u>Western Aphasia Bettery Aphasia Quotient</u> <u>Spatial index</u> - Composite score based on block design, facial recognition, judgment of line orientation <u>Right and left motor indexes</u> - Composite score based on grip strength and finger tapping	<u>LHS</u>: - n= 34 - With apraxia: n = 12 (limb apraxia errors: 6.0) - Without apraxia: n = 22 (limb apraxia errors: 1.5) - Age: mean = 61.0 yr - Poststroke: mean = 5.1 yr - Female: 21% - Aphasia quotient: 76.1 - Spatial index: 46.2 <u>RHS</u>: - n = 24 - Age: mean = 67.0 yr - Poststroke: mean = 3.9yr - Female: 42% - Aphasia quotient: 97.3 - Spatial index: 35.9 <u>LAB</u> (able-bodied, tested with left hand): - n = 41 - Female: 44% - Aphasia quotient: 98.9 - Spatial index: 50.2	**ANOVA (brain damage** [stroke, able-bodied] **× performing hand** [left, right]**)** <u>JHFT total time</u>: - Stroke groups (LHS, RHS) significantly more impaired than able-bodied groups (LAB, RAB) (LHS vs LAB: t = 3.8, $p < .001$; RHS vs RAB: t = 3.6, $p < .01$) - No significant interaction of brain damage by performing hand, indicating the similar impairments between stroke groups and able-bodied groups <u>JHFT writing time</u>: - Significant interaction of brain damage by performing hand (F = 4.4, $p < .05$) - LHS significantly more impaired than LAB (t = 3.5, $p < .01$), but no significant difference between RHS and RAB - Left hand performance significantly more impaired than right hand performance, indicating hand preference effect (F = 126.45, $p < .001$) **ANOVA (left brain damage × apraxia)** <u>JHFT total time</u>: - LHS with apraxia significantly more impaired than LHS without apraxia (t = 3.8, $p < .01$) and LAB (t = 7.5, $p < .001$) on the ipsilesional JHFT, but not on the ipsilesional motor index, indicating the importance of praxis for performing JHFT - No significant difference between LHS without apraxia and LAB <u>JHFT writing time</u>: - LHS with apraxia significantly more impaired than LHS without apraxia (t = 3.2, $p < .01$) and LAB (t = 6.6, $p < .001$) - No significant difference between LHS without apraxia and LAB **Stepwise linear regression** <u>JHFT total time</u>: - For LHS, severity of apraxia was the best predictor of JHFT total time: explaining 33% of the variance (F = 14.5, $p < .001$) - For RHS, the right (ipsilesional) motor index was the best predictor of JHFT total time: explaining 58% of the

Authors	Purpose of the study	Inclusion and exclusion criteria	Level of evidence	Independent variable	Dependent variable	Characteristics of subjects	Data analysis ----------------- Relevant findings
						RAB (able-bodied, tested with right hand): - n = 25 - Female: 40% - Aphasia quotient: 98.8 - Spatial index: 49.6	variance (F = 26.2, $p < .001$) - For RHS, the spatial index was a significant predictor of JHFT total time when only right motor index and spatial index were included as predictors: explaining 30% of the variance (F = 9.0, $p < .01$) - For RHS, the right motor index and spatial index together predicted JHFT total time: explaining 57% of the variance. There was a significant improvement in prediction when right motor index add into spatial index for predicting JHFT total time ($F_{change} = 12.9$, $p < .01$). JHFT writing time: - For LHS, severity of apraxia, not aphasia, was the best predictor of JHFT writing time: explaining 30 % of the variance (F = 12.4, $p < .01$), indicating apraxia is more important than language deficits for performing this task - For RHS, the right motor index was the best predictor of JHFT writing time: explaining 60 % of the variance (F = 28.7, $p < .001$) **Other results** - On the spatial index, LHS with apraxia significantly more impaired than LAB ($p < .05$), but no significant difference between LHS with and without apraxia - On the contralesional motor index, LHS with apraxia ($p < .001$) and without apraxia ($p < .01$) significantly more impaired than LAB - On the ipsilesional motor index, no significant impairments between LHS with apraxia and without apraxia and LAB, indicating the importance of praxis for performing JHFT

Note. LHS = Left hemispheric stroke. RHS = Right hemispheric stroke. JHFT = Jebsen-Taylor Hand Function Test. ADL = Activities of daily living.

Table 4-1 (Continued).

Authors	Purpose of the study	Inclusion and exclusion criteria	Level of evidence	Independent variable	Dependent variable	Characteristics of subjects	Data analysis ------------ Relevant findings
Yavuzer, G., Kucujdeveci, A., Arasil, T. & Elhan, A. (2001)	To identify the variables that best predict discharge functional status with the FIM	Inclusion: - First stroke - Admitted to inpatient rehabilitation - Infarction or Intracerebral hemorrhage Exclusion: - Subarachnoidal bleeding	**III**	Gender - Male - Female - Side of lesion - LHS - RHS Time - Admission - Discharge	FIM - Total - Gain - Cognitive	LHS: n = 39 RHS: n = 28 Whole sample: - Age: mean = 60.04 yr - Female: n = 43 - Male: n = 24 - Aphasia: 18.0% - Neglect: 21.0% - OAI: mean = 62.9 days - LOS: mean = 97.1 days	**Wilcoxon's matched pairs signed-rank test** *FIM scores at admission and discharge: the total group significantly improved from admission to discharge (p = .001) in ADL* **Pearson correlation** - Significant positive correlation between FIM admission and discharge scores ($r = .88$, $p < .001$) - Significant negative correlation between FIM admission and gain scores ($r = -.31$, $p < .05$) - Significant negative correlation between OAI and both FIM discharge and gain scores ($r = -.31$, $p < .01$; $r = -.29$, $p < .05$) **Student's *t* test** FIM scores at admission*: LHS were significantly less disabled than RHS (t = 2.15, p = .035)* - *RHS were significantly more impaired than LHS in FIM cognitive functions (p = .012)* **Multiple regression analysis** FIM discharge score: The OAI and the FIM admission score were the best predictors of FIM discharge scores ($R^2 = .80$), indicating patients with greater function at the time of admission and those with shorter OAI intervals would have greater function at the time of discharge

Note. FIM = Functional Independence Measure. LHS = Left hemispheric stroke. RHS = Right hemispheric stroke. OAI = Onset-admission interval. LOS = Length of stay.

Table 4-1 (Continued).

Authors	Purpose of the study	Inclusion and exclusion criteria	Level of evidence	Independent variable	Dependent variable	Characteristics of subjects	Data analysis --- Relevant findings
Chae, J. & Zorowitz, R. (1998)	"To examine the effects of cortical and subcortical infracts and lesion laterality on the functional status of stroke survivors" (p. 415)	<u>Inclusion:</u> - Medical record in stroke data bank of the National Institute of Neurological Disorders and Stroke - Admitted to acute inpatient rehabilitation during an 18-month period - Single non-hemorrhagic lesion <u>Exclusion:</u> - N/A	**III**	<u>Side of lesion</u> - LHS - RHS <u>Lesion level</u> - Cortical - Subcortical <u>Time</u> - Admission - Discharge	<u>FIM:</u> - Total - Domain o Self-care o Sphincter o Mobility o Locomotion o Communication o Social cognition - Gain	<u>Cortical + LHS</u>: - n = 19 - Age: mean = 64.4 yr - LOS: mean = 36.1 days - OAI: mean = 27.1 days - Male: 57.9% <u>Cortical + RHS</u>: - n = 21 - Age: mean = 67.6 yr - LOS: mean = 36.5 days - OAI: mean = 37.5 days - Male: 42.9% <u>Subcortical + LHS</u>: - n = 16 - age: mean = 67.9 - LOS: mean = 31.4 days - OAI: mean = 26.6 days - Male: 43.8% <u>Subcortical + RHS</u>: - n = 16 - Age: mean = 76.6 - LOS: mean = 31.1 days - OAI: mean = 29.7 days - Male: 56.3%	**MANOVA: (lesion level × hemisphere)** <u>FIM total score at discharge</u>: - Subcortical less disabled than cortical (significant lesion level main effect: F = 9.2, p < .01) - *RHS significantly less disabled than LHS (significant hemisphere level main effect: F = 5.5, p = .02)* <u>FIM self-care domain at admission</u>: - Subcortical less disabled than cortical (significant lesion level main effect: F = 8.2, p = .01), indicating the self-care domain depends, at some degree, on cognitive functions - *No significant difference between the LHS and RHS groups (hemisphere main effect: F = 1.3, p = .25), and RHS less disabled than LHS within cortical lesion, but not within subcortical lesion* <u>FIM sphincter function domain at admission</u>: *No significant difference between the LHS and RHS groups (hemisphere main effect: F = 2.2, p = .15), and RHS less disabled than LHS within cortical lesion, but not within subcortical lesion* <u>FIM mobility function domain at admission</u>: Subcortical less disabled than cortical (significant lesion level main effect: F = 5.2, p = .03), indicating the mobility domain depends, at some degree, on cognitive functions - *No significant difference between the LHS and RHS groups (hemisphere main effect: F = 0.3, p = .62), and LHS less disabled than RHS within cortical lesion, but not within subcortical lesion* <u>FIM locomotion function domain at admission</u>: *No significant difference between the LHS and RHS groups (hemisphere main effect: F = 0.2, p = .69), and LHS less disabled than RHS within cortical lesion, but not within subcortical lesion* <u>FIM communication domain at admission</u>: Significant interaction between lesion level and hemisphere: F = 5.3, p = .02), which may probably result from the aphasia associated with left hemisphere cortical lesions. <u>FIM communication and social cognition function domains at admission</u>: - Within LHS, subcortical less disabled than cortical (significant lesion level main effect: communication F = 11.3, p < .01; social cognition F = 5.4, p = .02)

Authors	Purpose of the study	Inclusion and exclusion criteria	Level of evidence	Independent variable	Dependent variable	Characteristics of subjects	Data analysis ----------- Relevant findings
							- Within cortical lesions, LHS more disabled than RHS (significant hemisphere main effect: communication F = 12.9, *p* < .01; social cognition F = 6.5, *p* = .01) FIM total score at discharge: - Subcortical less disabled than cortical (significant lesion level main effect: F = 7.8, *p* = .01) - *No significant difference between the LHS and RHS groups (hemisphere main effect: F = 3.1, p = .08), and RHS less disabled than LHS within cortical* FIM self-care domain at discharge: - Subcortical less disabled than cortical (significant lesion level main effect: F = 6.9, *p* = .01) - *No significant difference between the LHS and RHS groups (hemisphere main effect: F = 0.9, p = .34), and LHS less disabled than RHS within cortical lesion, but not within subcortical lesion* FIM sphincter function domain at discharge: *No significant difference between the LHS and RHS groups (hemisphere main effect: F = 0.1, p = .73), and RHS less disabled than LHS within cortical lesion, but not within subcortical lesion* FIM mobility function domain at discharge: *No significant difference between the LHS and RHS groups (hemisphere main effect: F = 0.6, p = .45), and LHS less disabled than RHS within cortical lesion, but equally disabled within subcortical lesion* FIM locomotion function domain at discharge: *No significant difference between the LHS and RHS groups (hemisphere main effect: F = 1.3, p = .25), and LHS less disabled than RHS within cortical and subcortical lesions* FIM communication and social cognition function domains at admission: - Within LHS, subcortical less disabled than cortical (significant lesion level main effect: Communication F = 9.1, *p* < .01; social cognition F = 11.4, *p* < .01) - Within RHS, cortical less disabled than subcortical (significant hemisphere main effect: Communication F = 7.6, *p* = .01; social cognition F = 6.4, *p* = .01) *FIM locomotion domain for gain score: LHS less disabled than RHS (significant hemisphere main effect: F = 3.9, p = .05)* *FIM communication domain for gain score: significant interaction between lesion level and hemisphere: F = 8.8, p < .01)*

Authors	Purpose of the study	Inclusion and exclusion criteria	Level of evidence	Independent variable	Dependent variable	Characteristics of subjects	Data analysis --------------- Relevant findings
							<u>FIM total score at admission and gain score</u>: *Patients with LHS at cortical level had lower admission FIM-total scores (FIM = 63.5) and higher FIM-total gains scores (FIM = 24.8), indicating they were more disabled at admission and showed more improvement in rehabilitation*

Note. LHS = Left hemispheric stroke. RHS = Right hemispheric stroke. FIM = Functional Independence Measure. LOS = Length of stay. OAI = Onset-admission interval.

Table 4-1 (Continued).

Authors	Purpose of the study	Inclusion and exclusion criteria	Level of evidence	Independent variable	Dependent variable	Characteristics of subjects	Data analysis ----------- Relevant findings
Ring, H., Feder, M., Schwartz, J. & Samuel, G. (1997)	To investigate the usefulness of the FIM total scores in measuring function at admission and discharge and functional gain during rehabilitation	Inclusion: - First stroke - Supratentorial stroke - Departmental database over 2-year period Exclusion: - Rehab unit does not accept patients with severe comorbidity (mainly cardiopulmonary or behavioral)	**III**	Side of lesion - LHS - RHS Syndrome - Aphasia in LHS - Neglect in RHS Lesion type - Infarction - Hemorrhage - Hemorrhagic infarction - No CT findings Lesion level - Cortical - Subcortical Time - Admission - Discharge	FIM - Total - Gain	LHS: - n = 67 - aphasia: 25.9% RHS: - n = 84 - neglect: 34.2% Whole sample: - Age: mean = 60.8 yr - Female: 40% - OAI: mean = 28.9 days - Rehabilitation: mean = 109.3 days - Thrombotic: 79.5% - Cortical: 69.1% - Subcortical: 29.9%	**Student's *t* test and ANOVA** FIM admission, discharge, and gain: - *For hemisphere main effects, no significant difference between FIM scores on admission and discharge* - For syndromes of neglect or aphasia after stroke, both syndrome groups were significantly more disabled than without syndrome groups on FIM scores at admission (neglect: $p = .001$; aphasia: $p = .01$) and discharge (neglect: $p = .02$; aphasia: $p = .09$), but they showed greater improvement in rehabilitation (neglect: $p = .09$; aphasia: $p = .04$) - Within lesion types, hemorrhagic lesion had the lowest FIM admission score (mean = 68.9) but the highest gain (mean = 35.6), indicating they were more disabled at admission but had the greatest improvement in rehabilitation - Within lesion sites, cortical lesion significantly more disabled on admission and discharge scores than subcortical lesion (admission: $p = .02$; discharge: $p = .03$), but they both achieved a similar FIM gain **Multiple linear regression** - For both LHS and RHS, the most significant predictors of functional gain (FIM gain score) were FIM admission score (LHS: $p = .002$; RHS: $p = .03$) and LOS (LHS: $p = .002$; RHS: $p = .011$), indicating rehabilitation after stroke is a time-dependent process - For both LHS and RHS, FIM admission scores related negatively to FIM gain scores (LHS: B= - .269; RHS: B = - .304), and LOS positively related to FIM gain scores (LHS: B= .208; RHS: B = .125)

Note. FIM = Functional Independence Measure. LHS = Left hemispheric stroke. RHS = Right hemispheric stroke. CT = computed tomography. OAI = Onset-admission interval. LOS = Length of stay.

Table 4-1 (Continued).

Authors	Purpose of the study	Inclusion and exclusion criteria	Level of evidence	Independent variable	Dependent variable	Characteristics of subjects	Data analysis ------------ Relevant findings
Johansson, B. B., Jadback, G., Norrving, B. & Widner, H. (1992)	To evaluate the medical and social outcome in patients with first stroke up to one year post-stroke	Inclusion: - First-ever stroke patients Exclusion: - Subarachnoidal hemorrhage	**III**	Side of lesion - LHS - RHS Time - Pre-stroke (before stroke onset) - Acute (less than 48hr after stroke onset) - 6m (after stroke onset) - 12m (after stroke onset) Age - Less than 75 years old - More than 75 years old	ADL Health Index (Pre, acute, 6m, 12m) o Dressing o Getting in and out of bed o Cutting hard food o Walking o Washing o Lavatory visits o Picking up things from the floor o Going out alone <QoL> NHP (high scores indicate more problems) (6m, 12m) o Mobility o Energy o Pain o Emotional reaction o Sleep o Social isolation Questionnaires, scales, and semi-structured interviews (12m) o Age o Personality o Coping o Diagnosis o Subjective symptoms o Attitude towards disease	LHS (Pre; ≤ 48 hrs): n = 155 RHS (Pre; ≤ 48 hrs): n = 134 LHS(12 months): n = 108 RHS(12 months): n = 96 Whole sample: - n = 346 - Female: 43.6% - Over 75 yr: 52.3% (n = 182) - Rehabilitation in acute wards: mean = 19 days - Rehabilitation at long-term hospital and nursing home: mean = 125.7 days	**Chi-square** ADL health index: - (Pre, 12m) In items of washing, going out alone ($p < .01$), get in and out of bed, and pick up things from the floor ($p < .05$), patients more than 75 years old were significantly more disabled than those less than 75 years old - (Pre, 12m) Many patients, particularly in the older age groups, were not totally independent in ADL before stroke onset - *(Pre, acute, 12m) RHS were consistently less independent than LHS over time on all items, indicating RHS may have a more serious prognosis than LHS in ADL, which may relate to neglect and reduced spatial awareness* - (Pre, acute, 6m, 12m) Many patients had complete functional independence 6m after an acute stroke Pain (Pre, 12m): - RHS had significantly more severe pain than LHS at both before and after stroke onset ($p < .05$), but neglect syndrome of RHS may bias the result **(forward/ backward) Multiple regression** - NHP (6m), coping interview (12m): The ability to wash oneself (taking a shower or a bath) was a highly significant ADL variable to predict QoL (no statistics available) - The most important predictors of QoL were ADL function, personality characteristics, awareness of disease and marital status **Step-wise discriminant analysis** QoL: - High QoL group (51% of 90 patients) managed nearly all ADL at 6m and 12m, had higher social activity before stroke (Pre), and had more possibilities to preserve and develop their feelings of identify and self value - *RHS may have a more serious prognosis than LHS with respect to life satisfaction, which may relate to neglect and reduced spatial awareness* **Other results** Language (Acute, 6m): - (Acute) 28% of 315 patients were considered to have dysphasia and 11% to have dysarthria - (Acute) 56 patients reported minor and 74 patients

Authors	Purpose of the study	Inclusion and exclusion criteria	Level of evidence	Independent variable	Dependent variable	Characteristics of subjects	Data analysis --------------- Relevant findings
					o Rehabilitation Pain (Pre, 6m, 12m) - Severe pain: required analgesic drugs every day - Moderate pain: frequent pain but not using drug every day Language (Acute, 6m) Type (by evaluation): - Dysphasia - Dysarthric Severity (by self-report): - Minor - Severe Social activity (Pre, 6m, 12m) - Visit friends (within the last 4 weeks) - Go to theater or any association or club (within the last 4 weeks) Help needed (Pre, 6m, 12m): for patients living in their homes, received from relatives/friends		reported severe language problem. - (6m) 28% of 259 patients reported language impairments and 8% of them graded severe language impairment, all but 3 patients with severe problems had LHS - (6m) Within patients who reported minor language impairments, 17 of them had lesions outside of LHS Social activity (6m, 12m): 77% of LHS and 81% of RHS reported low level of social activity. For RHS, the larger reduction in ADL and lower QoL did not result in more isolation Help needed (Pre): Many individuals needed much help before stroke and mortality was high in those with a pre-stroke handicap. Help needed (12m): Assistances from relatives and friends far exceeded those from community.

Note. ADL = Activities of daily living. QoL = Quality of life. NHP = Nottingham health profile. LHS = Left hemispheric stroke. RHS = Right hemispheric stroke.

Table 4-1 (Continued).

Authors	Purpose of the study	Inclusion and exclusion criteria	Level of evidence	Independent variable	Dependent variable	Characteristics of subjects	Data analysis ---------- Relevant findings
Shiotsuka, W., Burton, G. U., Pedretti, L. W. & Llorens, L. A. (1992)	To determine differences in the ADL performance between LHS and RHS	<u>Inclusion:</u> - A CVA of the cardiovascular system - Right hand dominant - Complete Klein-Bell Activities of daily living scores <u>Exclusion:</u> - Previous history of CVA - Bilateral hemispheric involvement - CVA of the basilar or cerebellar arteries - Amputation of a limb - Significant motor, perceptual, or cognitive dysfunction from a secondary diagnosis - Been discharged from the rehabilitation unit for a medical reason - Incomplete rehabilitation program	**III**	<u>Side of lesion</u> - LHS - RHS <u>Time</u> - Admission - Discharge	<u>Klein-Bell Activities of daily living</u> - Total - Item o Dressing o Elimination o Mobility o Bathing/hygiene o Eating o Emergency telephone communication <u>LOS</u>	<u>LHS</u>: - n = 42 - Female: n = 21 - Male: n = 21 - Age: mean = 67.57 yr - OAI: mean = 20.76 days - LOS: mean = 18.26 days <u>RHS</u>: - n = 53 - Female: n = 27 - Male: n = 36 - Age: mean = 71.9 yr - OAI: mean = 26.17 days - LOS: mean = 19.77 days	**ANOVA** *<u>Klein-Bell Activities of daily living at admission and discharge (total score or subtest score)</u>: No significant differences were found between LHS and RHS at admission or discharge for total or item scores* - *For all items scores at admission, LHS less disabled than RHS (no significant difference between the two groups).* - *For all items scores at discharge, RHS less disabled than LHS (no significant difference between the two groups).* **Other results** The lower the scores in Klein-Bell Activities of daily living (more disabled) at admission, the longer the LOS

Note. CVA = Cerebrovascular accident. ADL = Activities of daily living. LHS = Left hemispheric stroke. RHS = Right hemispheric stroke. LOS = Length of stay. OAI = Onset-admission interval.

Table 4-1 (Continued).

Authors	Purpose of the study	Inclusion and exclusion criteria	Level of evidence	Independent variable	Dependent variable	Characteristics of subjects	Data analysis --------------- Relevant findings
Bernspang, B. & Fisher, A. G. (1995)	- To differentiate the ability to perform ADL between patients with strokes and nondisabled controls - To determine the differences of abilities to perform ADL between LHS and RHS	<u>Inclusion:</u> - LHS or RHS in AMPS database - 60 years old or older - Performed 2 AMPS tasks - All subjects with stroke were currently receiving occupational therapy services during the time of evaluation <u>Exclusion:</u> N/A	**IV**	<u>Group</u> - LHS - RHS - Nondisabled control <u>Gender</u> - Male - Female	<u>AMPS</u> - Motor skill - Process skill	<u>LHS</u>: - n = 76 - Age: mean = 73.4 yr - Female: n = 38 - Male: n = 38 - Independent: n = 16 - Minimal assistance: n = 30 - Moderate/ Maximum assistance: n = 30 <u>RHS</u>: - n = 71 - Age: mean = 73.5 yr - Female: n = 37 - Male: n = 34 - Independent: n = 16 - Minimal assistance: n = 25 - Moderate/ Maximum assistance: n = 30 <u>Nondisable</u>: - n = 83 - Age: mean = 72.9 yr - Female: n = 42 - Male: n = 41 - Independent: n = 76 - Minimal assistance: n = 7	**ANOVA and post hoc (Tukey) test** <u>AMPS motor and process skills</u>: - *No significant difference between LHS and RHS* - Stroke patients (LHS and RHS) significantly more impaired than nondisabled controls ($p < .05$), which was more distinct in motor skills than in process skills **Many-faceted Rasch (MFR)** <u>AMPS motor skills</u>: - RHS and LHS differed significantly in specific motor skills (more than .25 logits of random variations in item calibrations) and RHS had greater motor impairments than LHS: RHS had difficulties with maintaining an even and appropriate pace, transporting task objects, and coordinating two body parts to stabilize task objects; LHS had difficulties with calibrating the appropriate force and extent of movement - 8% of nondisabled controls and 89% of stroke subjects had motor ability measures below the 2.0 logits cutoff - Many of stroke subjects with motor impairments remained able to live independently in the community, by compensating the motor impairments: 59% (19/32) of stroke subjects had motor skill ability measures below the 2.0 logits cutoff, but performed independently <u>AMPS process skills</u>: - No significant differences were found between RHS and LHS in process skills, indicating that the two groups were equally able to organize and adapt the actions of task performance. - 5% of normal subjects and 63% of stroke subjects had process skill ability measures below the 1.0 logits cutoff - Stroke subjects with low AMPS process skill were less independent and needed assistance to live in the community: 16% (5/32) of stroke subjects had a process skill ability measure below the 1.0 logits cutoff but performed independently

Note. AMPS = Assessment of motor and process skills. LHS = Left hemispheric stroke. IADL = Instrumental activities of daily living. RHS = Right hemispheric stroke.

Table 4-1 (Continued).

Authors	Purpose of the study	Inclusion and exclusion criteria	Level of evidence	Independent variable	Dependent variable	Characteristics of subjects	Data analysis ---------- Relevant findings
Heinemann, A. W., Linacre, J. M., Wright, B. D., Hamilton, B. B. & Granger, C (1993)	To scale the FIM with Rasch analysis To determine the similarity of scaled measures across impairment groups	<u>Inclusion:</u> - Rehabilitation inpatients who received services at hospitals [subscribing to the uniform data system (UDS) (more than 190 facilities subscribed to the UDS data management service)] <u>Exclusion:</u> - Admitted only for evaluation - Readmitted to rehabilitation	**IV**	<u>Impairment groups</u>: - LHS (right hemiparesis) - RHS (left hemiparesis) - Stroke-bilateral - Brain dysfunction - Neurologic condition - Spinal cord dysfunction - Amputation - Arthritis - Back pain - Orthopedic impairment - Cardiac impairment - Pulmonary impairment - Burns - Congenital deformity - Other	<u>FIM</u> - Items - Domains o Motor o Cognitive	13 impairments: - LHS: n = 5008 - RHS: n = 4745 - stroke-bilateral: n = 339 - brain dysfunction: n = 2427 - neurologic condition: n = 1394 - spinal cord dysfunction: n = 1727 - amputation: n = 1400 - arthritis: n = 1324 - back pain: n = 752 - orthopedic impairment: n = 6808 - cardiac impairment: n = 122 - pulmonary impairment: n = 162 - burns: n = 47 - congenital deformity: n = 38 - other: n = 1324 <u>Whole sample</u>: - Female: 53% - age: mean = 62.1 yr - OAI: mean = 113.2 days - LOS: mean =	**Rasch analysis** <u>FIM misfit:</u> - Excellent item fit for both motor and cognitive FIM measures (mean square within 3% of 1.0 (from .98 to 1.03) - Less fit motor items (good fit statistics for the scale as a whole): bowel and bladder management (reflecting clinicians rate both management and continence) and stair climbing (reflecting opportunity of evaluation) <u>FIM motor item difficulty</u>: - Feeding and grooming were easier - Stair climbing, tub/shower transfer, and locomotion were harder - *Items of feeding, grooming, bowel, and bladder were harder for patients with LHS than those with RHS.* - *Items of upper body dressing, bed transfers, toileting, toilet transfers, lower body dressing, and stairs were harder for patients with RHS than those with LHS* - *The difficulty levels in items of bathing, locomotion, and tub transfer seemed equal for patients with LHS and RHS* <u>FIM cognitive item difficulty</u>: - Comprehension and expression were easier - Problem solving was the most difficult **Factor analysis (results of Rasch analysis)** <u>FIM motor items</u>: - The first factor was defined as the relative difficulty of feeding to stair climbing and locomotion: stair climbing was always the most difficult item and feeding the easiest item - The second factor was defined as the difficulty of bladder and bowel management: this task was relatively easy for stroke survivors compared to other impairment groups <u>FIM cognitive items</u>: - The first factor was defined as the relative difficulty of verbal expression: Verbal expression was relatively difficult for RHS than LHS and bilateral stroke - The second factor was defined as the difficulty of social interaction: This task was relatively easy for LHS

Authors	Purpose of the study	Inclusion and exclusion criteria	Level of evidence	Independent variable	Dependent variable	Characteristics of subjects	Data analysis Relevant findings
						29.5 days	

Note. FIM = Functional Independence Measure. LHS = Left hemispheric stroke. RHS = Right hemispheric stroke. OAI = Onset-admission interval. LOS = Length of stay.

Table 4-1 (Continued).

Authors	Purpose of the study	Inclusion and exclusion criteria	Level of evidence	Independent variable	Dependent variable	Characteristics of subjects	Data analysis ----------- Relevant findings
Tsuji, T, Sonoda, S., Domen, K., Saitoh, E., Liu, M. & Chino, N. (1995)	To determine the difficulty pattern of the FIM in Japan using Rasch analysis To demonstrate the worldwide utility of the FIM by comparing the results obtained in the United States and Japan	<u>Inclusion:</u> - primary LHS or RHS - no history of peripheral n. injury or injury to the central nervous system <u>Exclusion:</u> - Cannot complete rehabilitation	**IV**	<u>Side of lesion</u> - LHS - RHS <u>Country</u> - America (data retrieved from Heinemann et al, 1993) - Japan	<u>FIM</u> - Items - Domains o Motor o Cognitive	<u>LHS</u>: n = 95 <u>RHS</u>: n = 95 <u>Whole sample</u>: - Age: 61.4 yr - Female: 38.94% - Cerebral infarction: n = 110 - Cerebral bleeding: n = 70 - Subarachnoid hemorrhage: n = 10 - OAI: mean = 47.3 days - LOS: mean = 90.9 days	**Rasch analysis** <u>FIM misfit</u>: stairs (less opportunities), bathing (narrow bathroom, patients prone to slip, too much help from caregivers), bowel and bladder management (rating both continence and management), and cognitive items (aphasia in LHS) **Rasch analysis (item difficulty)** <u>FIM motor item difficulty</u>: - *No difference was found between LHS and RHS* - Americans and Japanese patients performed similarly, but bowel and bladder management was easier for Japanese patients <u>FIM cognitive item difficulty:</u> - Expression was the easiest for RHS and most difficult for LHS - Similar performance between Americans and Japanese patients, but social interaction was easier for Japanese patients with LHS, because Japanese patients had few chances for group exercise or recreation than American patients.

Note. FIM = Functional Independence Measure. ADL = Activities of daily living. LHS = Left hemispheric stroke. RHS = Right hemispheric stroke. OAI = Onset-admission interval. LOS = Length of stay.

Table 4-1 (Continued).

Authors	Purpose of the study	Inclusion and exclusion criteria	Level of evidence	Independent variable	Dependent variable	Characteristics of subjects	Data analysis Relevant findings
Gardarsdottir, S. & Kaplan, S. (2002)	To test the construct validity of the A-ONE To determine differences in ADL performance and neurobehavioral impairment between LHS and RHS	Inclusion: - primary LHS or RHS - no history of peripheral nerve injury or injury to the central nervous system (previous CVA, head injury, or dementia) - at least 1 week post-CVA and medically stable - no previous OT intervention - consent to participate in the study Exclusion: - N/A	**V**	Side of lesion - LHS - RHS	A-ONE - FIS: ADL performance (dressing, grooming and hygiene, transfer, feeding, and communication domains) - NSIS: Impairment component	LHS: - n = 23 - Female: n = 10 - Male: n = 13 - Age: mean = 72 yr - OAI: mean = 23.09 days RHS: - n = 19 - Female: n = 8 - Male: n = 11 - Age: mean = 69 yr - OAI: mean = 24.37 days	**Mann Whitney *U* test** *FIS of A-ONE: RHS significantly less disabled than LHS on 3/18 performance tasks (shave/make-up: Z = -2.48, p = .013;* comprehension: Z = -2.81, p = .536; speech: Z = -3.70, p = .001*)* - *In dressing domain, LHS less disabled than RHS on items of shirt, shoes, and fastening, but RHS less disabled than LHS on socks item, and both groups were equal on pants item* - *In grooming and hygiene domain, LHS less disabled than RHS on items of wash face and brush teeth, but RHS less disabled than LHS on items of comb hair and shave/make-up* - *In transfers and mobility domain, LHS less disabled on items of sit up in bed and out of bed, but RHS less disabled on maneuver item* - *In feeding domain, RHS less disabled on all items (drink, finger feeding, use fork or spoon, and use knife)* - In communication domain, RHS less disabled on all items (comprehension and speech) NSIS of A-ONE: - LHS had more severe impairments of motor apraxia in dressing, grooming and hygiene, and feeding tasks; impairments of abnormal tone in grooming and hygiene tasks ($p < .05$) than RHS - RHS had more severe impairments of unilateral body neglect in dressing and grooming and hygiene tasks; impairments of abnormal tone in the dressing, transfers and mobility, and feeding tasks ($p < .05$) than LHS

Note. A-ONE = Arnadottir OT-ADL Neurobehavioral Evaluation. ADL = Activities of daily living. LHS = Left hemispheric stroke. RHS = Right hemispheric stroke. CVA = Cerebrovascular accident.

FIS = Functional Independence Scale of A-ONE. NSIS = Neurobehavioral Specific Impairment subscale of A-ONE. OAI = Onset-admission interval.

Table 4-1 (Continued).

Authors	Purpose of the study	Inclusion and exclusion criteria	Level of evidence	Independent variable	Dependent variable	Characteristics of subjects	Data analysis --- Relevant findings
Mills, V. & DiGenio, M. (1983)	To examine possible differences in functional abilities of patients with cerebrovascular accidents on the right or left side	Inclusion: - CVA - MCA - moderate motor involvement at admission - right handed - admitted to the rehabilitation hospital two to four weeks after onset of CVA Exclusion: - previous major neurological diagnosis - discharge for other medical problem or didn't complete the rehabilitation treatment program	**V**	Side of lesion - LHS - RHS	Mobility - Transfers - Ambulation ADL - Dressing - Upper extremity function Perception / information processing - Short-term memory - Judgment - Body scheme Language - Verbal expression - Auditory comprehension - Reading - Writing LOS	LHS: - n = 52 - Age: mean = 67.7 yr - Female: n = 33 - Male: n = 19 RHS: - n = 50 - Age: mean = 70.8 yr - Female: n = 30 - Male: n = 20	**Mann-Whitney *U* test** *Mobility, ADL, perception/information processing, and LOS: no significant difference was found between LHS and RHS, indicating patients with stroke had the same prognosis for functional rehabilitation regardless of the side of the lesion* Language: significant differences were found between LHS and RHS ($p < .001$) [no specific data available to identify which group performs better] **Spearman correlation** LOS: - Longer LOS was significantly correlated with more impaired language ($r = -.32$, $p < .01$) and more disabled ADL performance ($r = -.32$, $p < .01$) for LHS - Longer LOS was significantly correlated with more impaired mobility ($r = -.38$, $p < .003$) and more disabled ADL performance ($r = -.31$, $p < .02$) for RHS, which may result from the ability to sustain postures that were prerequisite for standing and ambulation Perception/information processing: - Less impaired perception/information processing was significantly correlated with less impaired language ($r = -.40$, $p < .005$) and less disabled ADL performance ($r = .32$, $p < .02$) for LHS - Less impaired perception/information processing was significantly correlated with less impaired mobility ($r = .28$, $p < .02$) for RHS ADL: - Less disabled ADL performance was significantly correlated with less impaired language ($r = .29$, $p < .02$) and mobility ($r = .43$, $p < .001$) for LHS, which both require sequencing skills - Less disabled ADL performance was significantly correlated with less impaired mobility ($r = .65$, $p < .001$) for RHS

Note. CVA = Cerebrovascular accident. MCA = middle cerebral artery. LHS = Left hemispheric stroke. RHS = Right hemispheric stroke. LOS = Length of stay.

4.1.6 Issues of scaling and Rasch analysis

Most ADL functional status measures, such as the Barthel Index (BI) and the Frenchay Activities Index (FAI) (Kenneth J. Ottenbacher, 1997; Rogers & Holm, 1998), utilize ordinal scales. The scores at the same scale levels may actually be different because the item distances along the scales are not even and because they may represent varied meanings (Merbitz et al., 1989). Linedeboom, Vermeulen, Holman, and Haan (2003) mentioned that traditional statistical methods that assume the data are on an interval scale could not solve the difficulties related to the ordinal structure of the item set for ADL assessments. Furthermore, the major limitation of the classical statistical methods cannot separate the examinee characteristics and test characteristics, which influences the interpretation of results and leads to difficulty when practitioners try to compare performance among different tests (Hambleton et al., 1991). Rasch analysis uses logarithmic transformation to convert ordinal data into interval data along a logit (log odds unit) scale (Bond & Fox, 2007). In this way, Rasch analysis can "use a single synthetic index to describe the full spectrum of functioning" for interpreting functional status (Barberger-Gateau, Rainville, Letenneur, & Dartigues, 2000, p. 310). Parameters of item difficulty and person ability are yielded from Rasch analysis so that the item and person information to be separated and interpreted at the same time using the same measurement units (Hays et al., 2000; Tesio, 2003). While several studies in the previous systematic review applied Rasch analyses, the methodologies of these studies (i.e., participants, assessments, and data analyses) may have biased the interpretation of results by anchoring the dataset to multiple diagnostic samples (Heinemann, Linacre, Wright, Hamilton, & Granger, 1993), or comparing groups using item difficulty rather than person ability (Tsiji et al., 1995). The methods of describing interval data of ADL performance in stroke survivors are still vague and so further investigation is required.

4.1.7 Hypotheses

The present study examined ADL disability in stroke survivors at 3 months post-stroke. Because clinical practitioners provide interventions related to specific task disabilities, instead of for domains of ADL or overall ADL, differences in performing ADL tasks at three levels (overall, domain, and task levels) between patients with LHS and RHS were explored and clarified by comparing estimations of task difficulty and person ability using Rasch analysis. Based on the findings from the literature review, we hypothesized that, at overall level, ADL performance (person) ability would be significantly greater for those with LHS compared to those with RHS. We also hypothesized that, at the domain level, the performance (person) ability of patients with LHS would also be significantly better than for those with RHS on the PASS cognitive-instrumental activities of daily living (CIADL) domain. However, there would be no difference in performance (person) ability among the two groups when performing PASS functional mobility (FM), personal care (PC), or physical instrumental activities of daily living (PIADL) domains. Because information on differences between the two groups when performing specific tasks was limited in the literature, in the current study we also explored relevant clinical differences at the task level.

4.2 METHODS

4.2.1 Participants

The participants for the current study were recruited in a large prospective stroke outcome study (IRB# 010593). The inclusion criteria were (a) admission to the University of Pittsburgh Medical Center (UPMC) (b) diagnosis of acute stroke or a transient ischemic attach (c) radiological data available from admission assessments (e.g., CT scan, MRI, and or Xenon quantitative cerebral blood flow [ExCT qCBF]) or pharmacological interventions (tPA), and (d) physician approval. Children were excluded from the study. No exclusion criteria based upon gender, race, ethnicity, or HIV status were applied.

4.2.2 Procedure

The team in the large prospective stroke outcome study admitted the participants with a primary diagnosis of acute stroke to UPMC Presbyterian Hospital, after eligibility requirements were verified. After obtaining informed consent from the patient or a proxy, the attending physician or nursing staff gathered and documented demographic data. The staff of the Department of Occupational Therapy administered the ADL assessments at 24 hours, 5 days, and 3, 6, 9, and 12-months post acute stroke event. The ADL assessments took place where the person was residing at the time if participants lived within a 150 mile radius of Pittsburgh. The staff who performed the assessment procedures were trained by the PI to achieve a minimum interobserver standard of >90% with the criterion assessor. The large prospective stroke outcome study was approved by University of Pittsburgh Institutional Review Board (IRB#010593) and followed the guidelines of confidentiality under HIPAA.

4.2.3 Instrument

The PASS is a criterion-referenced and performance-based instrument. There are 26 tasks, with 163 subtasks. Each task has subtasks which are the essential steps and the testing criteria for that task. The independence, safety, and adequacy to perform the tasks are evaluated by the PASS. Because the current study was delimited to the construct of independence, data from safety and adequacy were not included in the data analyses. The 26 tasks are further categorized into 4 domains, functional mobility (FM) (5 items, 28 subtasks), personal self-care (PC) (3 items, 26 subtasks), physical-instrumental activities of daily living (PIADL) (4 items, 22 subtasks), and cognitive-instrumental activities of daily living (CIADL) (14 items, 87 subtasks) (see Table 4-2). Each subtask has the independence score ranging from 0 (unable to perform task independently) to 3 (independent), based on the assistance level and frequency. The independence scores for each task were the average subtask scores within that task and were then recoded into integer scores for the Rasch analysis. During the assessment, assistance is provided only when needed, which is documented starting from the least assistive prompt to the most assistive and intrusive prompt. By averaging the subtasks scores for a task, the independence score for that task is obtained. Because the data for the oven-use task was not always available, data for the PASS included only 25 tasks in the current study. Validity of the PASS was referenced to common geriatric ADL/IADL instruments. Unidimensionality of the PASS independence construct was established by exploratory factor analysis, including 1158 subjects with depression, osteoarthritis, cardiopulmonary disease, and dementia, macular degeneration, stroke, and a cohort of well-elderly. The PASS tasks presented one dominant construct. Test-retest reliability on two consecutive days was r = 0.96 (Chisholm, 2005) (see Chapter 2 for a more complete description of the PASS).

Table 4-2: Performance Assessment of Self-care Skills (PASS) Tasks, by Domain

<u>Functional Mobility (FM)</u>
Bed transfers (move from prone to supine position and rise from bed)
Stair use (ascend and descend stairs)
Toilet transfers (sit and rise from a toilet)
Bathtub/shower transfers (enter and exit tub and/or shower)
Indoor walking (walk indoors)
<u>Personal Self-care (PC)</u>
Oral hygiene (clean teeth, dentures and/or mouth)
Trimming toenails (groom toenails)
Dressing (don and doff upper body and lower body clothing)
<u>Physical Instrumental Activities of Daily Living (PIADL)</u>
Carry garbage (bend, lift and carry garbage sack)
Changing bed linen (put on bed linens)
Sweeping (clean spillage on the floor using a broom and a dust pan)
Cleanup after meal preparation (perform clean up tasks after meal preparation)
<u>Cognitive Instrumental Activities of Daily Living (CIADL)</u>
Shopping (select and purchase grocery items)
Bill paying (write checks for sample utility bills)
Checkbook balancing (balance a checkbook after writing checks)
Mailing (prepare envelopes for mailing checks)
Telephone use (use telephone to obtain information)
Medication management (read medication information and organize medication according to prescription)
Obtaining information: auditory (obtain information from a radio announcement)
Obtaining information: visual (obtain information from a newspaper)
Small repairs (repair a flashlight)
Home safety (identify and correct hazards or problems in home safety situations)
Playing bingo (play bingo)
Oven use (cook muffins in an oven)
Stovetop use (cook soup on a stovetop)
Using sharp utensils (cut an apple with a sharp knife)

Note. This table (Chisholm, 2005, p. 10) is used with permission.

4.3 DATA ANALYSIS

4.3.1 Data preparation

The demographic and medical history data related to stroke were described by descriptive statistics, using SPSS version 14.0. Before performing Rasch analysis, missing data were replaced according to like subjects (i.e., LHS and RHS), using linear interpolation method in the SPSS (SPSS Inc., 2005). Rasch analysis was performed using WINSTEPS version 3.64.1.

4.3.2 Rasch analysis

4.3.2.1 Logits of item difficulty and person ability

Using Rasch analysis, the data from the PASS were transformed logits (log odds units). The item difficulty logits represent the underlying difficulty of an item in a measurement tool, whereas the person ability logits represent the underlying ability of the person performing all tasks in a tool.

4.3.2.2 Hierarchies of item difficulty and person ability

Rasch analysis yields hierarchies of item difficulty and person ability. With the PASS, items with more positive logit values were harder than those with more negative values, whereas the persons with more positive logit values had greater abilities to perform tasks independently than those with negative logit values. The current study set zero as the midpoint of difficulty level

(Bond & Fox, 2007; Linacre, 2007). To judge substaintial differences among item calibrations, the literature recommends the value of ±0.50 logits (Linacre, 1994; Silverstein et al., 1992).

4.3.2.3 Diagnosis: Construct validity

Fit statistics

The fit of the data to the Rasch Model are established with infit and outfit statistics of mean squares (mean of squared residuals). The recommended mean square error range for infit and outfit statistics, for clinical observations, is 0.5 to 1.7. When both values for a task item are out of this range (≤ 0.5 or ≥ 1.7), the item is a misfit, and requires further examination (Bond & Fox, 2007).

Principal component analysis of Rasch residuals

In principal component analysis (PCA) of Rasch residuals, the measurement construct is unidimensional when the explained variance of the model is greater than 60%, the unexplained variance of the first contrast is less than 5%, or the eigenvalue of the unexplained variance by the first contrast is smaller than 3.0. Additionally, factor sensitivity provides data about the influence of the minor trait(s), and the smaller the number, the better (Bond & Fox, 2007; Linacre, 2007).

Category function

Average measure, category fit statistics, thresholds, and category frequency are the factors to check category functions of a measurement tool. The model expects the average measures increase as the rating difficulties increase. For category fit statistics, the outfit mean squares are expected to be less than 2. Thresholds or category calibrations are also expected to increase with the category values. In category frequency, there should be at least 10 responses per category in the measurement tool for a sample.

Scale linearity

In the linearity plot, if the plots of the raw scores against the item difficulty logits, are closer to the ordinary least square regression line, the scale of a measurement tool is more linear. The scale linearity can be also determined by the R^2 linearity value.

4.3.2.4 Diagnosis: Reliability

Item reliability

The item reliability index confirms the stability of the item difficulty hierarchy and indicates that there is a wider range of item difficulty and a larger sample size (Bond & Fox, 2007; Linacre, 2007). The greater the value, the better the reliability of the tool.

Person reliability

A reliable tool is expected to have a person reliability index value higher than .80, indicating that the tool is reliable and the person ability ordering is stable. A person separation index over 2 implies the person stratification levels are stable for the measurement tool. A Cronbach's alpha value over 0.90 implies that the instrument is internally consistent (Bond & Fox, 2007; Linacre, 2007; Portney & Watkins, 2000).

4.3.2.5 Anchoring item values

Before analyzing a subset of data, the calibration of the items in an instrument should be anchored so that the item difficulty hierarchy is fixed. The person ability logits for a specific population should always be calculated based on the anchored values of the item difficulty logits (Bond & Fox, 2007).

4.3.3 Differential functioning

4.3.3.1 Differential group functioning (DGF): Overall PASS and PASS domains

Differential Group Functioning (DGF) was first introduced by Linacre (2007). It detects the differences of groups of items (latent classes) among groups of people. However, the interpretation of the DGF results were limited (Linacre, 2007) and there was a lack of information in previous literature. In the current study, the concept of DGF was applied to test our hypotheses regarding LHS versus RHS for the overall PASS and PASS domains. Because age is a factor known to impact stroke differentially (Johnston, Keith, & Hinderer, 1992) and gender is a factor required in DIF analysis, these two variables were considered potential covariates assuming they met the necessary assumptions required of ANCOVA. Pearson Product Moment correlations (age) and Point Biserial correlations (gender) were calculated using SPSS version 14.0 (Portney & Watkins, 2000; SPSS Inc., 2005).

To examine the differences at the overall PASS and domain levels, an independent t-test or an analysis of covariance (ANCOVA) was performed. In the analysis, the dependent variables were the person ability logits for the overall tasks in the PASS, or for the tasks in a PASS domain. The DGF were using SPSS version 14.0.

4.3.3.2 Differential item functioning (DIF): PASS tasks

Differential Item Functioning (DIF) in Rasch analysis is a method used to detect the differences among groups of people, or person classes in response to an item (e.g., LHS and RHS). To explore clinically relevant differences in task performance between participants with LHS and RHS, DIF was performed. Because the Rasch model holds to the principle that the items are invariant across testing contexts, the DIF exists when the difficulty logits of an item for the person classes differ more than the errors expected by the model. Some reasons that DIF is produced are relevant to clinical practice. That is, the item may be at its usual difficulty level for one person class (e.g., LHS), but it may be harder or easier for another person class (e.g., RHS). Alternative explanations are that a person class may perform a task at its usual ability level, but the other class may perform better or worse than the usual ability level. Sometimes the issue of fairness occurs when an item favors or disadvantages one person class over another (Bond & Fox, 2007; Linacre, 2007). In addition, to categorize the person classes in the current study DIF analysis also requires analysis of the impact of gender and age (Tennant & Conaghan, 2007). Gender was dichotomized as male/female, and age as above/below the median age of our sample (66 years old). DIF can also identify clinically relevant differences among person classes.

The Winsteps calculates the item difficulty logits for each person class, namely the DIF measure, for an item in a measurement tool (Linacre, 2007). For dichotomous groups of people, the DIF contrasts were calculated by subtracting the DIF measure from one person class from the other. A DIF contrast value over 0.50 logits, regardless the sign or direction, suggests potential DIF for an item. In the current study, substaintial differences in task performance between the two groups were identified if a DIF contrast value was over 0.50 logits. Furthermore, a student t-test is performed in Winsteps program to compare the DIF measures between two person classes for each item and probability corresponding to the t-value shows whether the difference appeared by chance. If the probability is less than 0.05 and the DIF contrast is more than 0.50 logits, the specific item has DIF, and needs to be examined further (Linacre, 2007).

4.4 RESULTS

4.4.1 Participants

The demographics of the participants in the anchored dataset are shown in Table 4-3. The sample was predominantly male (58.3%), white (97.6%), and with ischemic stroke (88.2%). The mean age of this sample was 63.24 years. Four types of strokes, left hemispheric stroke (51.2%), right hemispheric stroke (37.0%), cerebellum stroke (6.2%), and brainstem stroke (2.8%) were included in this sample. Other data from medical histories or factors related to stroke are also summarized in Table 4-3.

Table 4-4 presents the characteristics of the participants with LHS and RHS at 3 months post-stroke. The percentage of the two types of stroke in the sample was not significantly different ($\chi^2_{(1)} = 1.70$, $p = .19$). Both sample groups were predominantly male (60.98% for LHS and 63.33% for RHS) and white (95.1 for LHS and 100% for RHS). The mean age of LHS was 65.00 years and of LHS was 67.03 years. There was no significant difference in age between the two groups ($t_{(56)} = -0.61$, $p = 0.55$). For both groups, percentages of the participants with ischemic stroke were larger than those with hemorrhagic stroke, and significantly more RHS than LHS participants had a history of hypertension ($\chi^2_{(1)} = 5.08$, $p = .03$). Other data from medical histories or factors related to stroke are also summarized in Table 4-4.

Table 4-3: Participants' Demographics and Stroke Characteristics of the Anchoring Dataset (N = 211)

Gender	
Male (%)	58.3
Age, years (M, *SD*)	63.24 (15.26)
Ethnicity	
White (%)	97.6
Black (%)	1.9
Other (%)	0.5
Stroke location[a]	
Left Hemisphere (%)	51.2
Right Hemisphere (%)	37.0
Cerebellum (%)	6.2
Brainstem (%)	2.8
Stroke type	
Ischemic (%)	88.2
Hemorrhagic (%)	11.8
Has prior CVA (%)[b]	6.3
Has history of hypertension (%)[b]	57.3
Has history of cardiac medication use (%)[b]	35.4
Has history of diabetes mellitus (%)[b]	22.8
Current cigarette smoker (%)[b]	23.3
History of alcohol abuse (%)[b]	6.3
History of atrial fibrillation (%)[b]	13.6
Has used antiplates medication in week prior to stroke (%)[b]	29.1

Note. [a]n = 205 for stroke location. [b]n = 206 for each characteristic.

Table 4-4: Participants' Demographics and Stroke Characteristics of the 3 Month Post-Stroke Dataset (N = 71)

	LHS[a] (*N* = 41)	RHS[a] (*N* = 30)	*p* value
Gender			.84[d]
Male (%)	60.98	63.33	
Age, years (M, *SD*)	65.00 (15.08)	67.03 (12.12)	.55[b]
Ethnicity			.47[d]
White (%)	95.1	100.0	
Black (%)	2.4	0.0	
Other (%)	2.4	0.0	
Stroke type			.87[d]
Ischemic (%)	85.4	86.7	
Hemorragic (%)	14.6	13.3	
Has prior CVA (%)	7.3	10.7	.62[c, d]
Has history of hypertension (%)	56.1	82.1	.02[c, e]
Has history of cardiac medication use (%)	41.5	50.0	.48[c, d]
Has history of diabetes mellitus (%)	26.8	32.1	.63[c, d]
Current cigarette smoker (%)	24.4	21.4	.78[c, d]
History of alcohol abuse (%)	4.9	7.1	.69[c, d]
History of atrial fibrillation (%)	14.6	14.3	.96[c, d]
Has used antiplates medication in week prior to stroke (%)	31.7	42.9	.34[c, d]

Note. [a]No significant difference ($\chi^2_{(1)}$ = 1.70, p = .19) between the percentages of the participants with left versus right hemispheric stroke. [b]$t_{(56)}$ = - .61, p = .55. [c]n = 28 for each characteristic for RHS. [d]No significant differences among participants with LHS versus RHS. [e]Significant difference among participants with LHS versus RHS on history of hypertension ($\chi^2_{(1)}$ = 5.08, p = .02).
LHS = Left hemispheric stroke. RHS = Right hemispheric stroke.

4.4.2 Rasch analysis

4.4.2.1 Item Hierarchy of Performance Assessment of Self-Care Skills (PASS) tasks

The PASS task items were anchored using Rasch analysis, in which 211 administrations of the PASS were included in the analysis. The tasks of trimming toenails and changing bed linens were the most difficult tasks for the stroke population. In contrast, the tasks of bed transfer, oral hygiene, toilet use, indoor walking, and dressing were the easiest tasks (see Figure 4-1). Between tasks of dressing and indoor walking, a large gap of difference existed (±0.55 logits) (see Chapter 3 for a more complete description of item hierarchy results).

4.4.2.2 PASS diagnosis: Construct validity

The PASS construct of independence was unidimensional, as indicated by an explained variance of 91.7%, and the PCA of Rasch residuals. The overall categories of the PASS also functioned excellently, and the PASS scale was linear. Therefore, the PASS had excellent construct validity for independence (see Chapter 2 for a more complete description of construct validity diagnosis results).

4.4.2.3 PASS diagnosis: Reliability

Overall, the reliability of the PASS was excellent, with an item reliability index value of 0.98, a person separation index value of 2.84, and a Cronbach’s alpha of 0.96 (see Chapter 2 for a more complete description of reliability diagnosis results).

4.4.2.4 Person ability logits of the PASS for the two groups (LHS and RHS)

There were 71 participants at 3 months post-stroke who performed the PASS. Their person ability logits, derived from the anchored PASS item difficulty logits, ranged from -3.79 to 4.50. The mean for the total sample was 0.42, with a standard deviation of 1.73. Figure 4-1 presents the item-person map for the PASS. The item-person map illustrates the distributions of the abilities of the persons with LHS or RHS at 3 months post-stroke, along the ability logit scale, and the difficulties of the items from the PASS, along the difficulty logit scale. The distribution of the persons was clustered as were the items. Figure 4-2 further illustrates the distributions of each group. Overall, the patterns of the two groups were generally similar ($t_{(69)} = 0.83$, $p = 0.41$), but persons with LHS (mean = 0.57, SD = 1.73) had slightly higher person ability logits than those with RHS (mean = 0.22, SD = 1.74). Clinically, this implies that persons with LHS were slightly more independent than those with RHS.

On the item-person map, the mean of the LHS person ability logits centered around the item difficulty level of stovetop use, whereas the mean for the RHS person ability logits centered around sweeping (see Chapter 3 for a more complete description of item hierarchy results). This implies that participants with LHS, at 3 months post-stroke had a 50% probability of being totally independent on stovetop use and the probability of success on easier items below it on the hierarchy would likely be greater than 50%. Similarly, participants with RHS, at 3 months post-stroke, had 50% probability of being totally independent on sweeping and the probability of success on easier items below it on the hierarchy would be greater than 50%. Likewise, participants with LHS had less than a 50% probability of being independent on the items above stovetop use in the hierarchy (i.e., bending/lifting to take out the garbage, bill paying, check writing, mailing bills, changing bed linens, and trimming toenails). For those with RHS, they also had a 50% probability of not being independent on the items above sweeping in the hierarchy, which include all of the items listed for LHS plus cleaning up after meals, shopping, and use of sharp utensils.

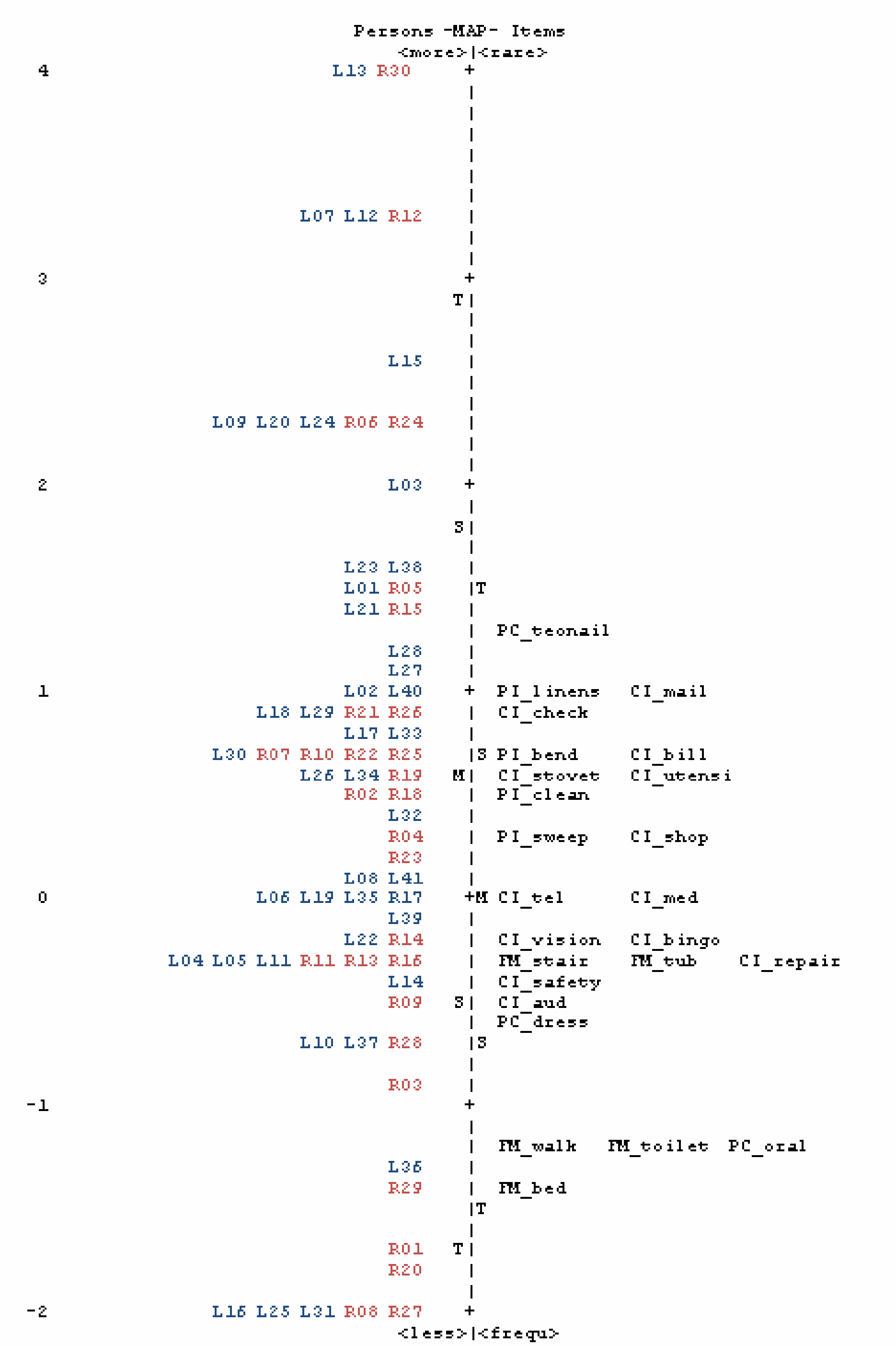

Figure 4-1: Item-person map for the Performance Assessment of Self-care Skills (PASS), with persons labeled side of stroke and .items labeled domains

To the left of the dashed line (the logit scale) is the person ability scale, distributed from the greatest ability at the top to the least ability at the bottom. Each letter/number represents a person in the sample. Numbers with an "L" in the front represent a person with left hemisphere stroke (LHS), whereas those with an "R" in the front represent person with right hemispheric stroke (RHS). To the right of the dashed line is the item difficulty scale distributed from the most difficult item at the top to the easiest item at the bottom. Item names with "FM_" in the front represent items in the functional mobility domain. Item names with "PC_" in the front represent items in the personal care domain. Item names with "PI_" in the front represent items in the physical instrumental activities of daily living domain. Item names with "CI_" in the front represent items in the cognitive instrumental activities of daily living domain. Along the logit scale, "M" represents the mean person ability or item difficulty estimate, "S" represents the location of one standard deviation (SD) from the mean estimates, and "T" is the second SD away from the mean estimates.

BEST PERFORMANCE (PERSON ABILITY)		
Logit values for LHS (*n* = 41)		Logit values for RHS (*n* = 30)
4.50	-4.50	4.50
	-4.00	
	-3.50	
3.31; 3.31		3.31
	-3.00	
2.65		
	-2.50	
2.04; 2.29; 2.29; 2.29		2.29; 2.29
	-2.00	
1.59; 1.59		
	-1.50	
1.10; 1.24; 1.40; 1.49		1.40; 1.49
	-1.00	
0.41; 0.58; 0.58; 0.72; 0.77; 0.77; 0.87; 0.93; 0.98; 0.98		0.63; 0.72; 0.72; 0.72; 0.72; 0.87; 0.93
	-0.50	0.50
0.04; 0.12; 0.12		0.24; 0.33; 0.54
0.00	-0.00	0.00
-0.04; -0.12; -0.16; -0.32; -0.32; -0.32; -0.45		-0.20; -0.28; -0.32; -0.32; -0.49
	0-0.50	
-0.67; -0.72		-0.67; -0.87
	0-1.00	
-1.29		-1.36
	0-1.50	
		-1.84; -1.72
	0-2.00	
	0-2.50	
	1-3.00	
	1-3.50	
-3.79; -3.79; -3.79		-3.79; -3.79
	1-4.00	
WORST PERFORMANCE (PERSON ABILITY)		

Note. LHS = Left hemispheric stroke. RHS = Right hemispheric stroke.

Figure 4-2: Rasch measures of person ability on the Performance Assessment of Self-care Skills (PASS) for participants with left and right hemispheric strokes (LHS and RHS).
Each number represents one person.

4.4.3 Differential functioning

4.4.3.1 Differential group functioning (DGF): Overall and domain levels

DGF was used to compare the independence of the participants with LHS and RHS in performing PASS tasks. The relationships between the potential covariates (age and gender) and dependent variables (person ability logits for each domain of the PASS) were checked (see Table

4-5). Age had significant relationships with the total PASS, and the PC and CIADL domains. While the significant relationships were not strong, age might be a clinically relevant factor that influences ADL performance. Therefore, age was included as a covariate in the ANCOVA for determining DGF (LHS, RHS) for the global PASS, PC domain, and CIADL domain. Independent t-tests were used to compare person ability logits differences for the FM and PIADL domains.

Table 4-5: Correlations among Person Ability Logits and Covariates (N = 71)

	PASS	FM	PC	PIADL	CIADL
Gender[a]	0.028	-0.032	-0.027	0.126	0.090
Age[b]	-0.292*	-0.146	-0.247*	-0.167	-0.336**

Note. PASS = Person ability logits for tasks in the Performance Assessment of Self-care Skills, including all 4 domains. FM = Person ability logits for tasks in functional mobility domain. PC = Person ability logits for tasks in person ability domain. PIADL = Person ability logits for tasks in physical instrumental activities of daily living domain. CIADL = Person ability logits for tasks in cognitive instrumental activities of daily living domain.
[a] Analyses were performed using Point Biserial correlations. [b] Analysis was performed using Pearson Product-Moment correlations.
* $p < .05$; ** $p < .001$

Overall PASS

An ANCOVA was used to examine independence of the participants with LHS and RHS in performing tasks of the global PASS. The assumption of homogeneity of variances was met, in which the Levene's test was not significant (p = 0.934). The assumption of homogeneity of regression was also met (F (1, 67) = 0.006, p = 0.937). The assumption of normality was not met for the LHS group (Shapiro-Wilk W = 0.093, p = 0.014 for LHS group, and Shapiro-Wilk W = 0.952, p = 0.186 for RHS group). All other assumptions (independence of the covariate and reliability of the covariate) were met. Age was significantly and negatively associated with ability to perform global PASS tasks (B = -0.038, F (1, 67) = 5.370, p = 0.024, partial η^2 = 0.074), for which 7.4% of the performance variance was explained by age. However, the ability of LHS and RHS participants to perform global PASS tasks was not significantly different (F (1, 67) = 0.003, p = 0.959), after adjusting for age (see Table 4-6). Finally, Figure 4-6 summarizes the statistically and clinically significant differences between participants with LHS and RHS at the overall level of the PASS.

PASS FM Domain

An independent t-test was used to examine participant independence in FM domain tasks. Participants with LHS were significantly more independent than those with RHS when performing PASS FM tasks ($t_{(69)}$ = 2.036, p = 0.046) (see Table 4-7). Finally, Figure 4-6 summarizes the statistically and clinically significant differences between participants with LHS and RHS at the FM domain level of the PASS.

PASS PC Domain

ANCOVA was used to examine participant independence in PC domain tasks. The assumption of homogeneity of variances was met, in which the Levene's test was not significant (p = 271). The assumption of homogeneity of regression was met (F (1, 67) = 0.006, p = 0.089). The assumption of normality was not met for either group (Shapiro-Wilk W = 0.873, p < 0.01 for LHS group, and Shapiro-Wilk W = 0.894, p = 0.006 for RHS group). All other assumptions were met. Age was not significantly associate with ability to perform PASS tasks in the PC domain (B = -0.024, F (1, 67) = 3.146, p = 0.081, partial η^2 = 0.045). The groups did not differ significantly on their ability to perform the PASS PC domain tasks (F (1, 67) = 0.269, p = 0.606) after adjusting for age (see Table 4-6). Finally, Figure 4-6 summarizes the statistically and clinically significant differences between participants with LHS and RHS at the PC domain level of the PASS.

PASS PIADL Domain

An independent t-test was used to examine participant independence in PIADL domain tasks. There was no significant difference between the two groups ($t_{(69)} = 1.325$, $p = 0.189$) (see Table 4-7). Finally, Figure 4-6 summarizes the statistically and clinically significant differences between participants with LHS and RHS at the PIADL domain level of the PASS.

PASS CIADL Domain

ANCOVA was used to examine independence of participants in CIADL domain tasks. The assumption of homogeneity of variances was met, in which the Levene's test was not significant ($p = 0.677$). The assumption of homogeneity of regression was also met ($F(1, 67) = 0.388$, $p = 0.535$). The assumption of normality was met for both groups (Shapiro-Wilk $W = 0.976$, $p = 0.528$ for LHS group, and Shapiro-Wilk $W = 0.945$, $p = 0.122$ for RHS group). All other assumptions were met. Being older (age) significantly and negatively impacted ability to perform PASS tasks in the CIADL domain ($B = -0.058$, $F(1, 67) = 0.388$, $p = 0.005$, partial $\eta^2 = 0.113$), in which age explained 11.3% of performance variance. PASS CIADL tasks were not significantly different between groups ($F(1, 67) = 0.262$, $p = 0.611$) after adjusting for age (see Table 4-6). Finally, Figure 4-6 summarizes the statistically and clinically significant differences between participants with LHS and RHS at the CIADL domain level of the PASS.

Table 4-6: Comparisons of Person Abilities of Participants with Left and Right Hemispheric Strokes (LHS and RHS) on the Performance Assessment of Self-care Skills (PASS) Domains, Using ANCOVA to Adjust for Age

Domains	Task items	M_L (n = 41)	M_R (n = 30)	F	p
PASS[a]		0.535	0.262	0.003	0.959
PC	Oral hygiene Dressing Trimming toenails	0.803	0.301	0.269	0.606
CIADL	Shopping Paying bills Checkbook balancing Mailing Telephone use Medication management Obtain information-auditory Obtain information-visual Small repairs Home safety Play bingo Stovetop use Use sharp utensils	0.602	0.391	0.262	0.611

Note. [a]PASS = Performance Assessment of Self-care Skills, including all 4 domains. PC = Personal care domain. PIADL = Physical-instrumental activities of daily living domain. M_L = Mean person ability logits of participants with LHS, after adjusting for age. M_R = Mean person ability logits of participants with RHS, after adjusting for age.

Table 4-7: Comparison of Person Abilities of Participants with Left and Right Hemispheric Strokes (LHS and RHS) on the Performance Assessment of Self-care Skills (PASS) Domains, Using Independent t-test

Domains	Task items	M_L (n = 41)	M_R (n = 30)	t	p
FM[b]	Bed transfers Stair use Toilet transfers Bathtub/Shower transfers Indoor walking	1.1339	0.1927	2.036	0.046*
PIADL[b]	Carrying garbage Changing bed linens Sweeping Clean up after meal preparation	0.7937	0.2987	1.325	0.189

Note. FM = Functional mobility domain. PIADL = Physical-instrumental activities of daily living domain. M_L = Mean person ability logits of participants with LHS. M_R = Mean person ability logits of participants with RHS.
* $p < .05$

4.4.3.2 Differential item functioning (DIF): Task level

DIF for side of stroke: Task level

DIF analysis was used to detect clinically relevant item difficulty differences for participants with LHS and RHS, at 3 months post-stroke. Tasks that were easier for participants with LHS were bed transfers, stair use, toilet transfers, oral hygiene, bathtub or shower transfers, dressing, shopping, carrying garbage, medication management, small repairs, sweeping, indoor walking, and home safety (see positive values of the DIF contrasts in Table 4-8).. Tasks that were easier for participants with RHS were trimming toenails, paying bills, checkbook balancing, mailing bills, telephone use, changing bed linens, obtaining auditory information, obtaining visual information, playing bingo, stovetop use, using sharp utensils, and cleanup after meal preparation (see negative values of the DIF contrasts in Table 4-8). Figure 4-3 visually illustrates the item difficulties for each group, and for each task. The tasks of stair use, toilet transfers, bathtub/shower transfers and oral hygiene had DIF contrasts over 0.50 logits, indicating substaintial differences, being easier for participants with LHS. Telephone use, obtaining auditory information and obtaining visual information had DIF contrasts over 0.50 logits, also indicating substaintial differences, being easier for participants with RHS. Among these tasks, telephone use, obtaining auditory information and obtaining visual information also showed differential item functioning (Prob. < 0.05), with all three tasks being significantly easier for participants with RHS. Finally, Figure 4-6 summarizes the statistically and clinically significant differences between participants with LHS and RHS at the task level of the PASS.

Table 4-8: Differential item functioning (DIF) analysis for interactions between tasks of PASS and groups of participants with LHS (n = 41) and RHS (n = 30).

PASS tasks	LHS DIF measure	RHS DIF measure	DIF contrast	t	Prob.
Bed transfers	-1.38	-0.92	0.46	1.06	0.0837
Stair use*	-0.75	-0.11	0.64	1.97	0.0535
Toilet transfers*	-1.84	-1.07	0.76	1.41	0.1623
Oral hygiene*	-1.51	-0.99	0.51	1.10	0.2734
Bathtub/Shower transfers*	-0.59	-0.06	0.53	1.69	0.0956
Trimming toenails	1.24	1.05	-0.19	0.29	0.5102
Dressing	-0.81	-0.36	0.45	1.33	0.1890
Shopping	0.29	0.55	0.26	0.95	0.3446
Paying bills	0.75	0.60	-0.16	-0.57	0.5726
Checkbook balancing	1.06	0.68	-0.38	-1.34	0.1836
Mailing	1.03	0.86	-0.17	-0.59	0.5593
Carrying garbage	0.42	0.81	0.39	1.40	0.1669
Telephone use**	0.29	-0.58	-0.87	-2.86	0.0058
Medication management	-0.09	-0.02	0.07	0.26	0.7957
Changing bed linens	0.96	0.86	-0.10	-0.34	0.7317
Obtaining information-auditory**	-0.24	-1.25	-1.00	-2.76	0.0077
Obtaining information-visual**	-0.05	-0.71	-0.65	-2.05	0.0443
Small repairs	-0.32	0.03	0.35	1.19	0.2378
Sweeping	0.05	0.34	0.29	1.03	0.3079
Indoor walking	-1.51	-1.25	0.25	0.53	0.5974
Home safety	-0.28	-0.21	0.07	0.25	0.8048
Playing bingo	-0.02	-0.21	-0.19	-0.64	0.5226
Stovetop use	0.69	0.64	-0.05	-0.17	0.8637
Using sharp utensils	0.79	0.68	-0.10	-0.38	0.7083
Cleanup after meal preparation	0.72	0.60	-0.12	-0.45	0.6562

Note. Modified matrix from the output of WINSTEPS. LHS = participants with left hemispheric stroke. RHS = participants with right hemispheric stroke. DIF measure =Item difficulty logits of the task in PASS for the group of participants (i.e., LHS or RHS). DIF contrast = Difference between the two DIF measures. t = DIF significance as a student's t-statistic (*d.f.* = 62). PASS = Performance Assessment of Self-care Skills. Prob. = Probability of observing of differences when there is no DIF.

* Statistically significant differences between groups of people (DIF contrasts > ± 0.50 logits).

** Differential item functioning items which also showed statistically significant differences between the two groups (DIF contrasts >± 0.50 logits and prob. < .05).

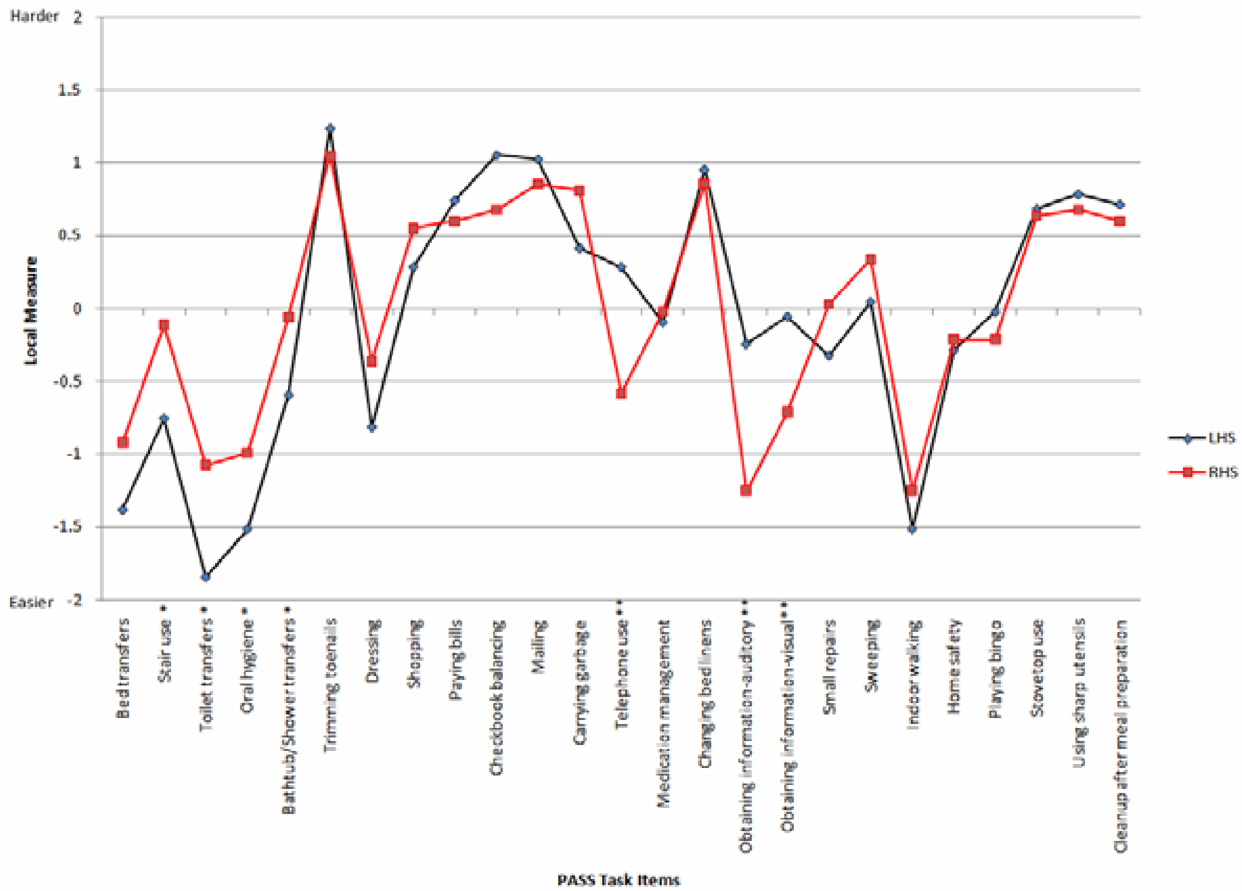

Figure 4-3: Differential item functioning (DIF) plot for side of stroke.
DIF plot on local measures, average item difficulty logits for each group (LHS and RHS), against task items of the Performance Assessment of Self-care Skills (PASS). LHS = Participants with left hemispheric stroke. RHS = Participants with right hemispheric stroke.
* Statistically significant differences between groups of people (DIF contrasts > ± 0.50 logits).
** Differential item functioning items which also showed statistically significant differences between the two groups (DIF contrasts >± 0.50 logits and prop. < .05).

DIF for gender: Task level

The DIF analysis was also used to detect the differences in item difficulty based on gender. Some tasks were easier for females than males (bed transfers, toilet transfers, oral hygiene, dressing, paying bills, telephone use, medication management, obtaining auditory information, small repairs, sweeping, indoor walking, stovetop use and cleanup after meal preparation) (see positive values of the DIF contrasts in Table 4-9), and some other tasks were easier for males than females (stair use, bathtub or shower transfers, trimming toenails, shopping, checkbook balancing, carrying garbage, changing bed linens, obtaining visual information, home safety, playing bingo, and using sharp utensils) (see negative values of the DIF contrasts in Table 4-9). The difficulty level of the mailing task for males and females was equal (DIF contrast = 0.00). Figure 4-4 visually illustrates the item difficulties for each group, and for each task. The tasks of bed transfers, dressing, sweeping, and cleanup after meal preparation were easier for females and had DIF contrasts over 0.50 logits, indicating substaintial differences, and potential DIF items. The tasks of trimming toenails, obtaining visual information, and playing bingo, were easier for males and had DIF contrasts over 0.50 logits, indicating substaintial differences, and potential DIF items. Among these tasks, trimming toenails and sweeping showed differential item functioning (Prob. < 0.05), with trimming toenails being significantly harder for females, and sweeping being significantly harder for males.

Table 4-9: Differential item functioning (DIF) analysis for interactions between tasks of PASS and groups of male (n = 44) and female participants (n = 27).

PASS tasks	Male DIF measure	Female DIF measure	DIF contrast	t	Prob.
Bed transfers*	-0.92	-1.53	0.60	1.38	0.1737
Stair use	-0.52	-0.24	-0.28	-0.82	0.5774
Toilet transfers	-1.25	-1.53	0.27	0.58	0.5628
Oral hygiene	-1.18	-1.27	0.09	0.21	0.8308
Bathtub/Shower transfers	-0.45	-0.10	-0.35	-1.07	0.2873
Trimming toenails**	0.90	1.74	-0.84	-2.64	0.0104
Dressing*	-0.42	-1.04	0.63	1.66	0.1018
Shopping	0.39	0.44	-0.05	-0.18	0.8577
Paying bills	0.71	0.63	0.07	0.25	0.8040
Checkbook balancing	0.85	1.01	-0.16	-0.53	0.5958
Mailing	0.96	0.96	0.00	0.00	1.0000
Carrying garbage	0.58	0.63	-0.06	-0.19	0.8463
Telephone use	-0.01	-0.17	0.15	0.47	0.6388
Medication management	-0.01	-0.17	0.15	0.47	0.6388
Changing bed linens	0.87	1.01	-0.13	-0.44	0.6610
Obtaining information-auditory	-0.60	-0.65	0.05	0.15	0.8806
Obtaining information-visual*	-0.48	0.05	-0.53	-1.62	0.1094
Small repairs	-0.16	-0.17	0.01	0.02	0.9817
Sweeping**	0.36	-0.31	0.68	2.10	0.0402
Indoor walking	-1.25	-1.53	0.27	0.58	0.5628
Home safety	-0.25	-0.24	-0.01	-0.03	0.9790
Playing bingo*	-0.31	0.38	-0.69	-2.22	0.0301
Stovetop use	0.73	0.51	0.23	0.75	0.4552
Using sharp utensils	0.74	0.76	-0.02	-0.07	0.9412
Cleanup after meal preparation*	0.82	0.32	0.05	1.63	0.1072

Note. Modified matrix from the output of WINSTEPS. DIF measure =Item difficulty logits of the task in PASS for the group of participants (i.e., male or female). DIF contrast = Difference between the two DIF measures. t = DIF significance as a student's t-statistic (*d.f.* = 62). PASS = Performance Assessment of Self-care Skills. Prob. = Probability of the reported t, indicating the probability of observing of these data when there is no DIF.

* Statistically significant differences between groups of people (DIF contrasts > ± 0.50 logits).

** Differential item functioning items which also showed statistically significant differences between the two groups (DIF contrasts >± 0.50 logits and prob. < .05).

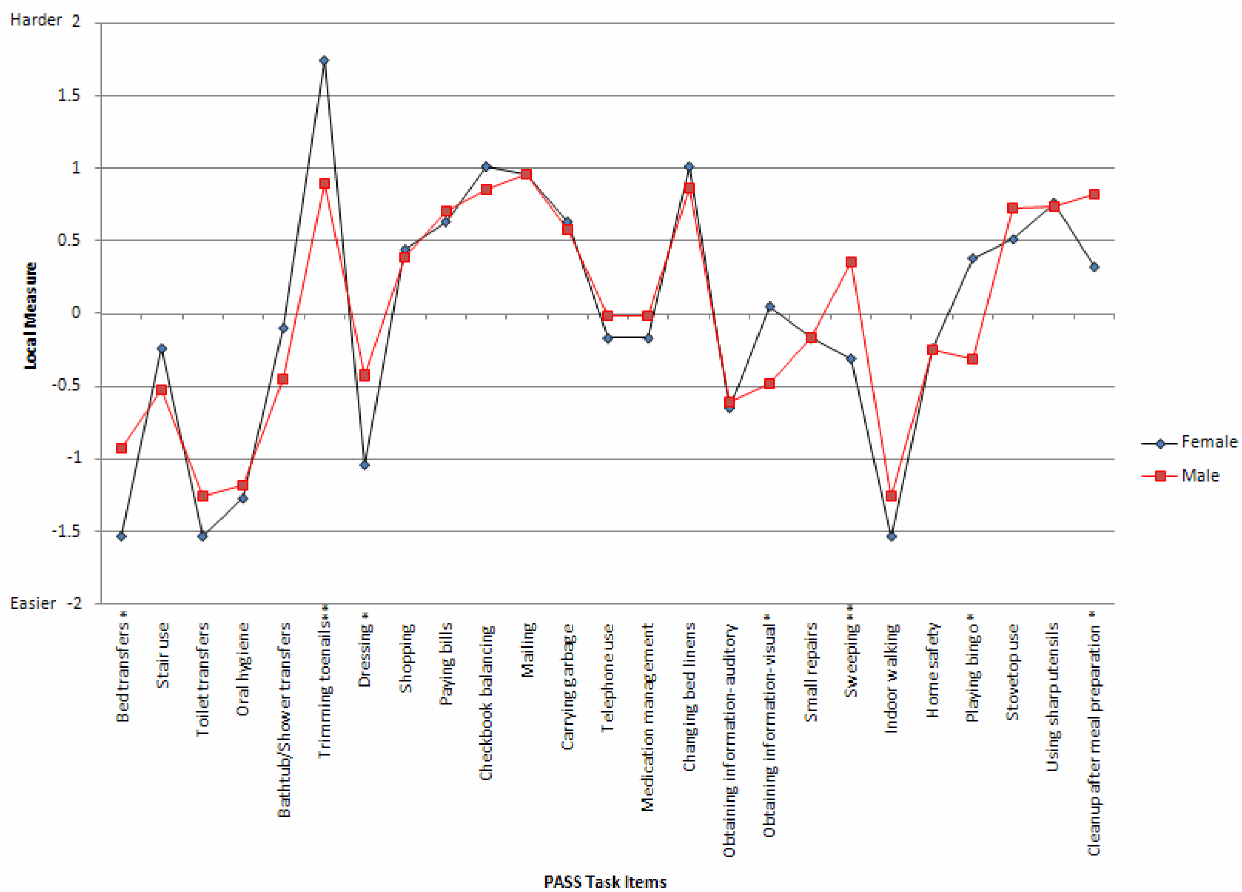

Figure 4-4: Differential item functioning (DIF) plot for gender.
DIF plot on local measures, average item difficulty logits for each group (male and female), against task items of the Performance Assessment of Self-care Skills (PASS).
* Statistically significant differences between groups of people (DIF contrasts > ± 0.50 logits).
** Differential item functioning items which also showed statistically significant differences between the two groups (DIF contrasts >± 0.50 logits and prop. < .05).

DIF for age: Task level

DIF analysis was then used to detect differences in item difficulty based on age. Some tasks were easier for younger participants (< 66 years) than for older participants (trimming toenails, shopping, paying bills, checkbook balancing, carrying garbage, telephone use, medication management, changing bed linens, obtaining visual information, small repairs, home safety, and playing bingo) (see positive values of the DIF contrasts in Table 4-10). Likewise, some tasks were easier for older participants (> 66 years) (bed transfers, stair use, toilet transfers, oral hygiene, bathtub or shower transfers, dressing, obtaining auditory information, small repairs, indoor walking, stovetop use, using sharp utensils, and cleanup after meal preparation) (see negative values of the DIF contrasts in Table 4-10). The difficulty level of the mailing task for both groups was equal (DIF contrast = 0.00). Figure 4-5 visually illustrates the item difficulties for each group, and for each task. The tasks of bed transfers and oral hygiene were more difficult for younger participants and had DIF contrasts over 0.50 logits, indicating there were substaintial differences, and potential DIF. However, DIF did not exist for age with the PASS.

Table 4-10: Differential item functioning (DIF) analysis for interactions between tasks of PASS and groups of younger (n = 31) and older (n = 40) participants.

PASS tasks	Young DIF measure	Old DIF measure	DIF contrast	t	Prob.
Bed transfers*	-0.76	-1.46	-0.70	-1.71	0.0914
Stair use	-0.24	-0.60	-0.36	-1.15	0.2540
Toilet transfers	-1.20	-1.46	-0.26	-0.56	0.5742
Oral hygiene*	-0.92	-1.46	-0.54	-1.28	0.2060
Bathtub/Shower transfers	-0.20	-0.46	-0.27	-0.88	0.3827
Trimming toenails	0.87	1.42	0.55	1.86	0.0677
Dressing	-0.45	-0.71	-0.26	-0.79	0.4341
Shopping	0.28	0.51	0.23	0.82	0.4136
Paying bills	0.65	0.71	0.06	0.23	0.8208
Checkbook balancing	0.70	1.06	0.37	0.31	0.1957
Mailing	0.96	0.96	0.00	0.00	1.0000
Carrying garbage	0.41	0.75	0.34	1.24	0.2206
Telephone use	-0.19	0.05	0.25	0.85	0.3975
Medication management	-0.15	0.02	0.17	0.58	0.5659
Changing bed linens	0.91	0.92	0.01	0.05	0.4773
Obtaining information-auditory	-0.52	-0.71	-0.19	-0.56	0.5760
Obtaining information-visual	-0.44	-0.21	0.23	0.76	0.4510
Small repairs	-0.39	0.01	0.41	1.35	0.1805
Sweeping	0.40	-0.01	-0.42	-1.49	0.1416
Indoor walking	-1.22	-1.46	-0.24	-0.53	0.6006
Home safety	-0.41	-0.14	0.27	0.90	0.3714
Playing bingo	-0.11	-0.10	0.01	0.03	0.9768
Stovetop use	0.73	0.61	-0.12	-0.44	0.6644
Using sharp utensils	0.86	0.64	-0.22	-0.79	0.4329
Cleanup after meal preparation	0.86	0.51	-0.35	-1.25	0.2174

Note. Modified matrix from the output of WINSTEPS. Young = Participant's age was younger than and equal to 66 years old. Old = Participant's age was older than 66 years old. DIF measure =Item difficulty logits of the task in PASS for the group of participants (i.e., young or old). DIF contrast = Difference between the two DIF measures. t = DIF significance as a student's t-statistic (*d.f.* = 62). PASS = Performance Assessment of Self-care Skills. Prob. = Probability of the reported t, indicating the probability of observing of these data when there is no DIF.
* Statistically significant differences between groups of people (DIF contrasts > ±0.50 logits)

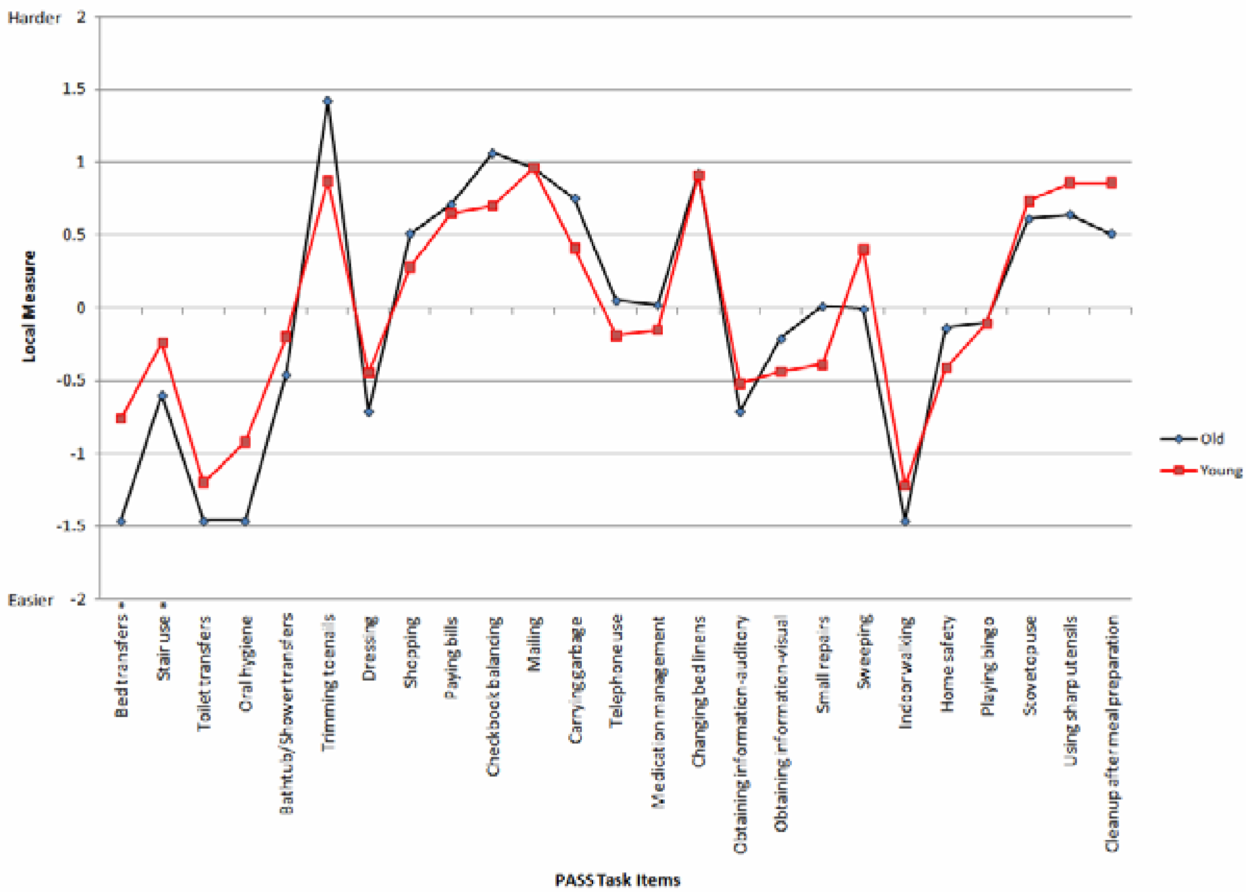

Figure 4-5: Differential item functioning (DIF) plot for age.
DIF plot on local measures, average item difficulty logits for each group (young and old), against task items of the Performance Assessment of Self-care Skills (PASS). Young = Participant's age was younger than and equal to 66 years old. Old = Participant's age was older than 66 years old.
* Statistically significant differences between groups of people (DIF contrasts > ±0.50 logits)

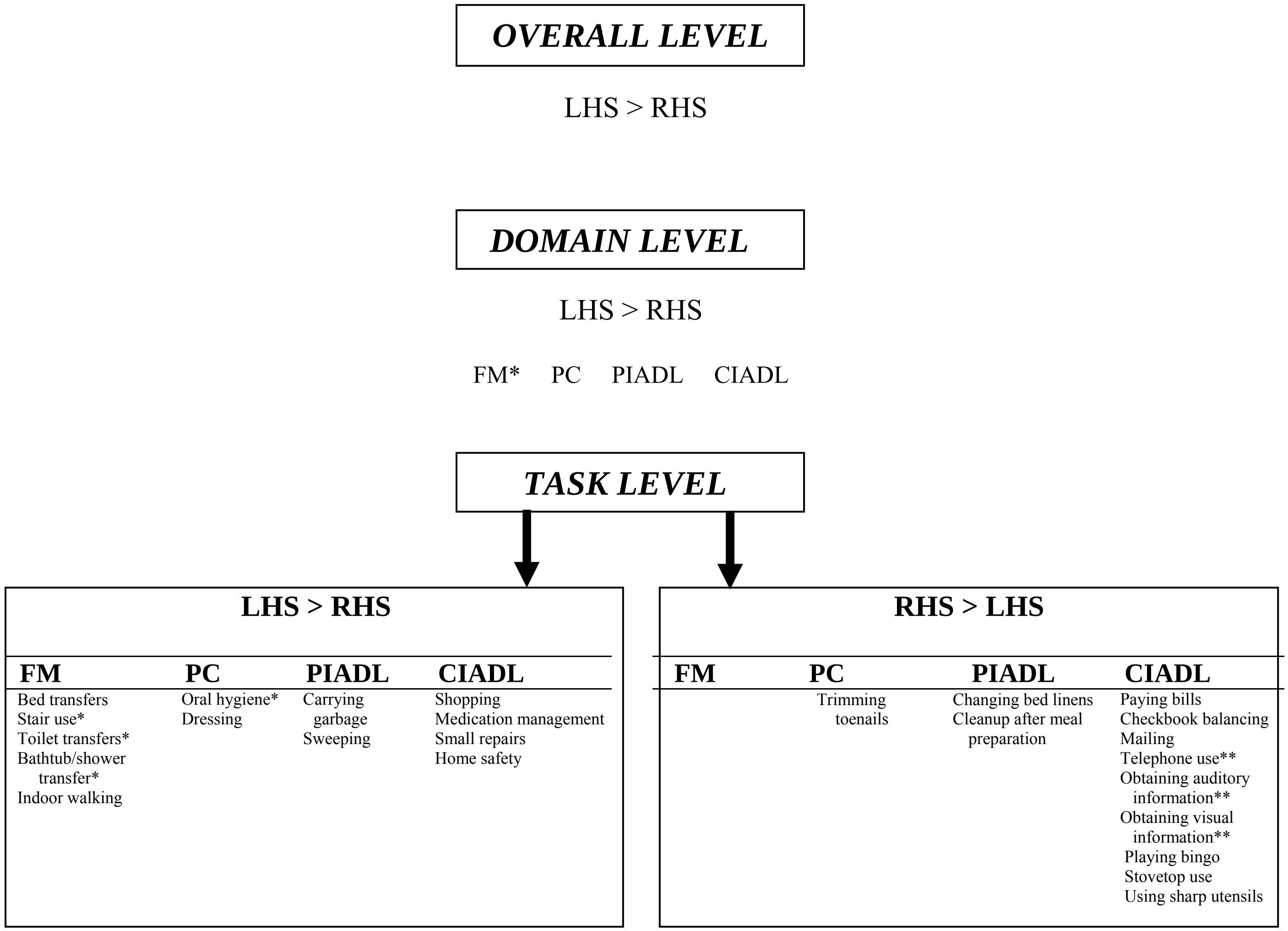

Figure 4-6: Summary of all statistically and clinically significant differences between participants with LHS and RHS at the overall, domain and task levels of the Performance Assessment of Self-Care Skills (PASS).

FM = functional mobility. PC = personal activities of daily living. PIADL = physical instrumental activities of daily living. CIADL = cognitive instrumental activities of daily living. > = performance was more independent in one group than the other, or task was easier for one group than the other.
* = Statistically significant or substaintially different ($p < .05$ in DGF, & DIF contrasts >± 0.50 logits in DIF). ** = DIF items (DIF contrasts >± 0.50 logits and probability < .05)

4.5 DISCUSSION

The purpose of the current study was to describe the differences in disabilities, at 3 months post-stroke, between persons with LHS and RHS, and to compare the impact of LHS and RHS on performance of overall ADL, ADL domains, and specific ADL tasks. Findings from our data did not support our hypothesis that participants with LHS were significantly more independent than those with RHS at the overall ADL level. Our hypotheses that the two groups would not differ in independence for PC and PIADL domains was supported. However, our hypothesis that there would be no difference between the groups for the FM domain was not supported. Participants with LHS were significantly more independent that those with RHS when for the PASS FM domain. Likewise, findings from our data did not support our hypothesis that participants with LHS were significantly more independent than those with RHS for the CIADL domain. At the task level, Rasch analysis was useful for identifying an item hierarchy of harder and easier tasks for each group, as well as clinically relevant differences in independence among FM, PC, CIADL and PIADL tasks between the two groups

Consistent with our findings of no significant differences between participants with LHS and RHS at the overall ADL level of independence at 3 months post-stroke, other studies have also often found the two groups to be equally independent at admission to rehabilitation (Ring et al., 1997; Shiotsuka et al., 1992), or discharge from rehabilitation (Chae & Zorowitz, 1998; Mills & DiGenio, 1983; Ring et al., 1997), however, differences at 3 months post-stroke have not been reported. Gardarsdottir and Kaplan (2002) explained that the similar performance between the two groups was not unexpected because both hemispheres contribute important and necessary functions when patients perform ADL tasks and that task performance relies on the combined central nervous system activity of both hemispheres. In contrast to our findings, Chae and Zorowitz (1998) found that those with RHS were significantly more independent, whereas Yavuzer et al. (2001) found that those with LHS were significantly more independent, but this was at admission to rehabilitation. Johansson et al. (1992) also found that those with LHS were consistently, but not significantly, more independent on the ADL Health Index pre-stroke, during the acute phase and at the 12 month follow-up, which was also consistent with our findings. Also consistent with the Johansson et al. (1992) study, we found that being older was significantly related to being less independent in overall ADL performance. Johansson et al. (1992) showed that the patients more than 75 years old were significantly less independent than those younger than 75 years, although in their sample persons 75 years and older had often been dependent in ADL performance before their strokes. However, in contrast to Johansson et al., our sample was considerably younger (mean age of 66 years), and when we adjusted for age, there was no significant difference in overall ADL independence between participants with LHS and RHS.

Our finding that participants with LHS were significantly more independent on the FM items (bed transfer, toilet transfer, tub/shower transfer, walk indoors, and stairs) at 3 months post-stroke was unexpected. However, this is somewhat consistent with the findings of Chae and Zorowitz (1998) who found that participants with LHS improved significantly more from admission to discharge than did those with RHS on two FIM locomotion items (push/walk, stairs). Our findings are also in contrast to the Rasch analysis by Heinemann et al. (1993) who found that during rehabilitation participants with RHS were more independent on FIM bed transfer, toilet transfer, and stair items. Two of the five PASS FM items from the current study overlapped with the Chae and Zorowitz study (walk, stairs), and three PASS items matched the items in the Heinemann et al. study (bed transfer, toilet transfer, and stairs). However, neither study measured tub/shower transfers, nor did they assess all five PASS FM items, as in the current study, which may account for the difference in findings.

The current study was seminal because the task performance of the stroke survivors was examined using a criterion-referenced, performance based assessment, and so the data were more objective than those reported in previous literature. Moreover, the PASS extended into the IADLs. In addition, no literature was found that discussed clinically relevant differences between the two groups for ADL tasks. The item-person map showed that patients with LHS at 3 months post-stroke were generally more able than those with RHS, but still had difficulty with some IADL tasks (e.g., stovetop use, financial tasks, and changing bed linens) at the more difficult level of the hierarchy. Those with RHS also had difficulty with the IADL tasks, but the person ability level of many participants with RHS was centered in the PC and FM items. Furthermore, our data showed that some PASS tasks (i.e., stair use, toilet transfers, bathtub/shower transfers, oral hygiene) were clinically and significantly more difficult for the patients with RHS than for those with LHS. In the Heinemann et al. (1993) study, using Rasch analysis, the LHS group was more independent than RHS on toilet transfers in the FIM, which was in contrast to our finding.

Our data also showed that some PASS tasks (i.e., telephone use, obtaining auditory information and obtaining visual information) showed differential item functioning, in which these tasks were significantly more difficult for the patients with LHS than for those with RHS. Telephone use was also more difficult for LHS subjects in the study by Shiotsuka et al. (1992) study (emergency telephone communication at discharge). For those tasks with substaintial differences between the two group, we assumed that the underlying neurological impairments of LHS (e.g., aphasia) and RHS (e.g., neglect) contributed to greater difficulty performing a specific task (Johansson et al., 1992). However, current thinking in rehabilitation suggests that searching for the differences in task and subtask performance between the two groups is more practical, and may also shed light on the underlying impairment mechanism contributing to the problems in performance (Skidmore, 2003).

Differences between genders in the performance of ADL tasks also showed clinically relevant differences. Clinically relevant differences in task performance that were easier for males usually required more physical demands (e.g., trimming toenails, carrying garbage, stair use, bathtub/shower transfers) and some financial tasks (e.g., shopping, checkbook balancing). Only trimming toenails was statistically significantly easier for males. Clinically relevant differences that were easier for females were usually related to home management tasks and personal care (e.g., meal preparation, sweeping, and dressing). Only sweeping was substaintially easier for females. Gender differences in task performance have been reported in the literature, including differences among stroke survivors, with males performing better on physical tasks, and females performing better on household tasks (Appelros, 2007). In the study by Appelros (2007), on the IADL items of the Frenchay Activities Index (FAI), females were more independent than males on light housework, and males were more independent on actively pursuing hobbies. However, whether these differences were clinically significant was not clear, and FAI is not a performance-based or criterion-referenced assessment. Our findings regarding gender differences suggested that practitioners need to be aware of gender differences in task performance among stroke survivors, especially if the person lives alone, and must perform all tasks independently (i.e., males must prepare meals; females must carry out the garbage).

The findings from the current study showed that no DIF item existed for age, indicating that the items do not function differently based on age alone (Bond & Fox, 2007). However, the DGF showed that age was significantly associated with independence in performing CIADL tasks. Further investigation revealed that most of the CIADL tasks (e.g., financial tasks, medication management, telephone use, obtaining critical information) were harder for older stroke survivors, regardless of side of stroke. Based on this finding, practitioners should remain vigilant during the performance of the CIADL tasks, especially if the client with stroke is older (e.g., over 66 years old), and plans to return to independent living in the community.

Limitations and Recommendations

As with all studies, our study had limitations. Our sample size was not large enough to compare more specific focal stroke groups. For example, the study by Chae and Zorowitz (1998) examined performance of stroke groups based on cortical or subcortical lesions. Because our sample also included participants with both ischemic and hemorrhagic strokes, there may be greater variance in some of the test items. In addition, traditional method of dealing with those items exist, and following DIF such items are often eliminated and the data reanalyzed. In the current study, we kept these items in the analyses because the DIF could be explained by our clinical experiences. Moreover, the current study was an exploratory study and the DIF phenomenon requires further evaluation in future studies. Future studies should seek to use more homogeneous samples, and compare performance of stroke survivors with LHS and RHS on PASS item subtasks. Rasch analysis of subtask performance could further delineate critical differences between the groups at a finer level of detail. In addition, exploring the differences in independence for stroke survivors at different time points (e.g., beyond 3 months post-stroke) is recommended so that patterns of stroke performance progression can be documented, especially for the IADL.

4.6 CONCLUSION

The current study compared differences in independence on 25 ADL tasks between stroke survivors with LHS and RHS, at 3 months post-stroke. The findings from the criterion-referenced and performance-based PASS showed no differences between participants with LHS and RHS on the overall PASS, nor on the PC, CIADL, and IADL domains. However, those with LHS were significantly more independent on the FM domain than those with RHS. Furthermore, the current study was the first to identify clinically relevant differences in specific ADL tasks between groups based on hemisphere of stroke, age, and gender Of clinical relevance, the item-person map showed that, overall, the person abilities of the LHS group indicated greater overall ability than those in the RHS group, and the map also delineated which items on the item difficulty hierarchy were more difficult for members of each group. These findings will help to guide practitioners in identifying less and more difficult tasks to focus on during intervention. DIF Rasch analysis revealed that at the item level, participants with LHS were more independent in 13 PASS items, and RHS were more independent in 12 items. Rasch DIF analysis also revealed several differences in task performance between genders, and age differences on the CIADL tasks, indicating that practitioners need to take these factors into consideration when assessing functional status. Further study was recommended to investigate the differences of the two groups at the subtask level so that the specific aspects of task breakdown between stroke survivors at 3 months post-stroke could be delineated more clearly for clinical practice.

5.0 CONCLUSION

The purpose of this dissertation was to describe disability in ADL tasks of stroke survivors at 3 months post-stroke, from multiple perspectives. The general aims of this study were to:

(1) confirm, correlate, and compare the measurement properties of five functional status instruments that are used often with stroke survivors, and use different scoring systems to rate ADL disability,

(2) compare specific disability items for stroke survivors at 3-months post-stroke, and

(3) describe the differences in disabilities between persons with LHS and RHS and compare the functional impact of LHS and RHS on specific ADL tasks.

The first study examined the measurement constructs of five measures (the Glasgow Outcome Scale, 5-point version [GOS5], the Glasgow Outcome Scale, 8-point version [GOS8], the Modified Rankin Scale [mRS], the Barthel Index [BI], and the Performance Assessment of Self-Care Skills [PASS]), and then correlated the scores from each tool to understand the relationship among the five measures. Finally, the items from the five measures were converted to a single metric to compare item difficulty for independence, as well as the interaction between item difficulty and person ability. The Rasch analysis Partial Credit Model (PCM) showed that the five measures evaluated the independence of the stroke survivors somewhat differently, although the relationships of the scores from the five measures were significant and moderate to strong. Our findings indicated that data from the ADL measures using different scoring systems and assessment methods should be interpreted carefully because some instruments (e.g., GOS5, GOS8, and mRS) also involve other concept/constructs in addition to independence. Rasch analysis diagnostics also indicated that the tools are not interchangeable.

The second study extracted out the global non-summative measures (GOS5, GOS8, and mRS) to compare the difficulty of the items in the rehabilitation-relevant BI and PASS tools in depth. BI and the PASS items were converted into a single metric, using the Rasch analysis. Partial Credit Model (PCM), PCM, and common person equating showed that overall, the PASS overall was more difficult for the stroke survivors at 3 months post-stroke. Examining differences in the difficulty of common BADL items in these two assessments showed that the PASS task items were always more difficult than the BI items. In addition, the item-person map of the BI-PASS equating measure revealed that the participants were challenged by IADL tasks in the performance-based PASS, and overestimated their ability in the informant-based BI (BADL tasks). These findings indicate that practitioners should be aware of the need to assess IADL task performance of stroke survivors who live in the community, and that the item hierarchy of the equating measure could also be used to guide evaluation and intervention priorities.

The third study began with a systematic review of the literature comparing differences in ADL independence between stroke survivors with LHS and RHS. The literature lacked information about differences in independence between LHS and RHS groups on specific tasks, especially IADL tasks. The study then examined the differences of the two groups at the overall ADL, ADL domain, and specific ADL task levels of the PASS. Rasch analysis and differential group functioning (DGF) revealed that there were no significant differences between the independence of the two groups at the overall PASS level, as well as for the PC, PIADL, and

CIADL domains. However, the LHS group performed significantly more independently on items in the FM domain. Rasch analysis and differential item functioning (DIF) revealed that participants with LHS were more independent on 13 PASS items, and participants with RHS were more independent on 12 PASS items. The impact of gender and age on task performance was also examined, and there were differences based on both factors. Based on this finding, the practitioners should consider these factors for those items during evaluations. The item-person map provides detailed information about differences in task difficulty for persons with LHS and RHS, and can help to guide practitioners as they select easier or more difficult tasks for interventions with stroke survivors who live in the community.

In summary, the current study described levels of independence in performing specific ADL tasks in stroke survivors at 3 months post-stroke, as well as how the construct of an ADL measure could influence clinical decisions by practitioners. In particular, Rasch analysis and its family models yielded data that were clinically relevant to stroke survivors and the practitioners who treat them.

BIBLIOGRAPHY

American Heart Association. (2007). Heart and Stroke Statistical Update - 2007. *Circulation, 115*, e69-e171.

Anemaet, W. K. (2002). Using standardized measures to meet the challenge of stroke assessment. *Topics in Geriatric Rehabilitation 18*(2), 47-62.

Appelros, P. (2007). Characteristics of the Frenchay Activities Index one year after stroke: A population-based study. *Disability and Rehabilitation, 29*, 785-790.

Ashburn, A. (1997). Physical recovery following stroke. *Physiotherapy, 83*, 480-490.

Bernspång, B., & Fisher, A. (1995). Differences between persons with right or left cerebral vascular accident on the assessment of motor and process skills. *Archives of Physical and Medical Rehabilitation, 76*, 1144-1151.

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences.* (2nd ed.). Mahwah, NJ: Lawrence Erlabaum.

Brandstater, M. E. (1990). An overview of stroke rehabilitation. *Stroke, 21* (Suppl II), II-40-II-42.

Caelo, A. D., Lamassa, M., Baldereschi, M., Pracucci, G., Basile, A. M., Wolfe, C. D. A., et al. (2003). Sex differences in the clinical presentation, resource use, and 3-month outcome of acute stroke in Europe: Data from a multicenter multinational hospital-based registry. *Stroke, 34*, 1114-1119.

Carod-Artal, F. J., Gonzalez-Gutierrez, J. L., Herrero, J. A. E., Horan, T., & de Seijas, E. V. (2002). Functional recovery and instrumental activities of daily living: Follow-up 1-year after treatment in a stroke unit. *Brain Injury, 16*, 207-216.

Carter, L. T., Oliveira, D. O., Duponte, J., & Lynch, S. V. (1998). The relationship of cognitive skills performance to activities of daily living in stroke patients. *American Journal of Occupational Therapy, 42*, 449-455.

Chae, J., & Richard, Z. (1998). Functional status of cortical and subcortical nonhemorrhagic stroke survivors and the effect of lesion laterality. *American Journal of Physical Medicine & Rehabilitation, 77*, 415-420.

Chisholm, D. (2005). *Disability in older adults with depression.* Unpublished doctoral dissertation, University of Pittsburgh, Pittsburgh, PA.

Dallmeijer, A. J., Dekker, J., Roorda, L. D., Knol, D. L., van Baalen, B., de Groot, V., et al. (2005). Differential item functioning of the functional independence measure in higher performing neurological patients. *Journal of Rehabilitation Medicine, 37*, 346-352.

Dittmar, S. S. (1997). Overview: A functional approach to measurement of rehabilitation outcomes. In S. S. Dittmar & G. E. Gresham (Eds.), *Functional assessment and outcome measure for the rehabilitation health professional* (pp. 1-10). Gaithersburg, MD: Aspen Publishers, Inc.

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. *Archives of Physical and Medical Rehabilitation, 84*, 950-963.

Duncan, P. W., Jorgensen, H. S., & Wade, D. T. (2000). Outcome measures in acute stroke trials: A systematic review and some recommendations to improve practice. *Stroke, 31*, 1429-1438.

Duncan, P. W., Zorowitz, R., Bates, B., Choi, J. Y., Glasberg, J. J., Graham, G. D., et al. (2005). Management of Adult Stroke Rehabilitation Care: A Clinical Practice Guideline *Stroke, 36*, e100-e143.

Ferrucci, L., Bandinelli, S., Guralnik, J. M., Lamponi, M., Bertini, C., Falchini, M., et al. (1993). Recovery of functional status after stroke: A postrehabilitation follow-up study. *Stroke, 24*, 200-205.

Fillenbaum, G. G. (1988). *Multidimensional functional assessment of older adults: The Duke older Americans resources and services procedures*. Hillsdale, NJ: Lawrence Erlbaum.

Finlayson, M., Havens, B., Holm, M. B., & Denend, T. V. (2003). Integrating a performance-based observation measure of functional status into a population-based longitudinal study of aging. *Canadian Journal on Aging, 22*, 185-195.

Fisher, Jr., W. (1992). Reliability Statistics. *Rasch Measurement Transactions, 6*, 238.

Forerch, C., Misselwitz, B., Sitzer, M., Berger, K., Steinmetz, H., & Neumann-Haefelin, N. (2005). Difference in recognition of right and left hemispheric stroke. *The Lancet, 366*, 392-393.

Gardarsdóttir, S., & Kaplan, S. (2002). Validity of Arnadottir OT-ADL neurobehavioral evaluation (A-ONE): Performance in activities of daily living and neurobehavioral impairments of persons with left and right hemisphere damage. *The American Journal of Occupational Therapy, 56*, 499-508.

Granger, C. V., Dewis, L. S., Peters, N. C., Sherwood, C. C., & Barrett, J. E. (1979). Stroke rehabilitation: Analysis of repeated Barthel Index measures. *Archives of Physical and Medical Rehabilitation, 60*, 14-17.

Granger, C. V., Hamilton, B. B., & Graesham, G. E. (1988). The stroke rehabilitation outcome study - Part I: General description. *Archives of Physical and Medical Rehabilitation, 69*, 506-509.

Gurland, B., Kuriansky, J., Sharpe, L., Simon, R., Stiller, P., & Birkett, P. (1977). The Comprehensive Assess,emt and Referral Evaluation (CARE). *International Journal of Aging and Human Development, 8*, 9-42.

Hafsteinsdóttir, T. B., & Grypdonck, M. (1997). Being a stroke patient: A review of the literature. *Journal of Advanced Nursing, 26*, 580-588.

Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. Newbury Park, CA: Sage.

Harrington, D. L., & Haaland, K. Y. (1992). Motor sequencing with left hemisphere damage: are some cognitive deficits specific to limb apraxia? *Brain, 115*, 857-874.

Hays, R. S., Morales, L. S., & Reise, S. P. (2000). Item response theory and health outcomes measurements in the 21th century. *Medical Care, 38*(9, Suppl. II), II-28-II-42.

Heinemann, A. W., Linacre, J. M., Wright, B. D., Hamilton, B. B., & Granger, C. (1993). Relationships between impairment and physical disability as measured by the Functional Independence Measure. *Archives of Physical and Medical Rehabilitation, 74*, 566-573.

Holm, M. B., & Rogers, J. C. (1999). Performance Assessment of Self-Care Skills. In B. J. Hemphill-Pearson (Ed.), *Assessments in occupational therapy mental health.* (pp. 117-124). Thorofare, NJ: Slack.

Hsueh, I.-P., Wang, W.-C., Sheu, C.-F., & Hsieh, C.-L. (2004). Rasch analysis of combining two indices to assess comprehensive ADL function in stroke patients. *Stroke, 35*, 721-726.

Jennett, B., & Bond, M. (1975). Assessment of outcome after severe brain damage: A practical scale. *Lancet, 1*, 480-484.

Jennett, B., Snoek, J., Bond, M., & Brooks, N. (1981). Disability after severe head injury: observations on the use of the Glasgow Outcome Scale. *Journal of Neurology, Neurosurgery, & Psychiatry, 44*, 285-293.

Johansson, B. B., Jadbäck, G., Norrving, B., & Widner, H. (1992). Evaluation of long-term functional status in first-ever stroke patients in a defined population. *Scandinavian Journal of Rehabilitation Medicine, Suppl. 26*, 105-114.

Johnston, M. V., Keith, R. A., & Hinderer, S. E. (1992). Measurement standards for interdisciplinary medical rehabilitation. *Archives of Physical and Medical Rehabilitation, 73*(12-S), S3-S23.

Kasner, S. E. (2006). Clinical interpretation and use of stroke scales. *Lancet Neurology, 5*, 603-612.

Katz, S., Ford, A. B., Moskowitz, R. W., Jackson, B. A., & Jaffe, M. W. (1963). Studies of illness in the aged. The Index of ADL: A standardized measure of biological and psychosocial function. *Journal of the American Medical Association, 185*, 914-919.

Lai, S. M., & Duncan, P. W. (2001). Stroke recovery profile and the modified rankin assessment. *Neuroepidemiology, 20*, 26-30.

Lawton, M. P. (1983). Environment and other determinants of well-being in older people. *Gerontologist, 23*, 349-357.

Lawton, M. P., & Brody, E. M. (1969). Assessment of older people: self-maintaining and instrumental activities of daily living. *Gerontologist, 9*, 179-186.

Linacre, J. M. (1994). Sample size and item calibration stability. *Rasch Measurement Transactions, 8*, 328.

Linacre, J. M. (1997). KR-20/Cronbach alpha or Rasch reliability: Which tells the "truth"? *Rasch Measurement Transactions, 11*, 580-581.

Linacre, J. M. (2007). *A user's guide to WINSTEPS.* (3.64.1 ed.). Chicago: MESA press.

Lind, K. (1982). A synthesis of studies on stroke rehabilitation. *Journal of Chronical Diseases, 35*, 133-149.

Lindeboom, R., Vermeulen, M., Holman, E., & Haan, R. J. (2003). Activities of daily living instruments: Optimizing scales for neurologic assessments. *Neurology, 60*, 738-742.

Lundgren-Nilsson, A., Grimby, G., Ring, H., Tesio, L., Lawton, G., Slade, A., et al. (2005). Cross-cultural validity of functional independence measure items in stroke: A study using Rasch analysis. *Journal of Rehabilitation Medicine, 37*, 23-31.

Lyden, P., Lu, M., Jackson, C., Marler, J., Kothari, R., Brott, T., et al. (1999). Underlying structure of the National Institutes of Health Stroke Scale: Results of a factor analysis. *Stroke, 30*, 2347-2354.

Merbitz, C., Morris, J., & Grip, J. C. (1989). Ordinal scale and foundations of misinference. *Archives of Physical and Medical Rehabilitation, 70*, 308-312.

Mills, V. M., & DiGenio, M. (1983). Functional differences in patients with left or right cerebrovascular accidents. *Physical Therapy, 63*, 481-487.

Moskowitz, E., & McCann, C. B. (1957). Classification of disability in the chronically ill and ageing. *Journal of Chronic Diseases, 5*, 342-346.

Nagi, S. Z. (1965). Some conceptual issues in disability and rehabilitation. In M. B. Sussman (Ed.), *Sociology and rehabilitation* (pp. 100-113). Washington DC: American Sociological Association.

Nuutinen, J., Liu, Y., Laaksa, M. P., Karonene, J. O., Roivainen, R., Vanninen, R. L., et al. (2006). Assessing the outcome of stroke: a comparison between MRI and clinical stroke scale. *Acta Neurologica Scandinavica, 113*, 100-107.

Ottenbacher, K. J. (1997). Methodological issues in measurement of functional status and rehabilitation outcomes. In S. S. Dittmar & G. E. Gresham (Eds.), *Functional assessment and outcome measure for the rehabilitation health professional* (pp. 17-26). Gaithersburg, Maryland: Aspen Publishers, Inc.

Ottenbacher, K. J., Hsu, Y., Granger, C. V., & Fiedler, R. C. (1996). The reliability of the functional independence measure: A quantitative review. *Archives of Physical and Medical Rehabilitation, 77*, 1226-1232.

Pfeffer, R. I., Kurosaki, T. T., Karrah, C. H., Chance, J. M., & Filos, S. (1982). Measurement of functional activities in older adults in the community. *Journal of Gerontology, 37*, 323-329.

Portney, L. G., & Watkins, M. P. (2000). *Foundations of clinical research: Application to practice.* (2nd ed.). Upper Saddle River, NJ: Prentice-Hall, Inc.

Raina, K. D. (2005). *Disability in order women with heart failure.* Unpublished doctoral dissertation, University of Pittsburgh, Pittsburgh, PA.

Rankin, J. (1957). Cerebral vascular accidents in patients over the age of 60. *Scottish Medical Journal, 2*, 200-215.

Ravaud, J.-F., Delcey, M., & Yelnik, A. (1999). Construct validity of the functional independence measure (FIM): Questioning the unidimensionality of the scale and the "value" of FIM scores. *Scandinavian Journal of Rehabilitation Medicine, 31*, 31-41.

Reuben, D. B., & Solomon, D. H. (1989). Assessment in geriatrics of caveats and names. *Journal of American Geriatric Society, 37*, 570-572.

Ring, H., Feder, M., Schwartz, J., & Samuels, G. (1997). Functional measures of first-stroke rehabilitation independence measure total score with a clinical rationale. *Archives of Physical and Medical Rehabilitation, 78*, 630-635.

Rogers, J. C. (1983). Eleanor Clarke Slagle Lectureship - 1983; Clinical reasoning: The ethics, science, and art. *American Journal of Occupational Therapy, 37*, 601-616.

Rogers, J. C., & Holm, M. B. (1998). Evaluation of occupational performance areas. Section 1: Evaluation of activities of daily living and home management. In M. E. Neistadt & E. B. Crepeau (Eds.), *Willard & Spackman's Occupational Therapy* (9th ed., pp. 185-208). Philadelphia: Lippincott Williams & Wilkins.

Rogers, J. C., Holm, M. B., Beach, S., Schulz, R., Cipriani, J., Fox, A., et al. (2003). Concordance of four methods of disability assessment using perofrmance in the home as the criterion method. *Arthritis & Rheumatism, 49*, 640-647.

Rogers, J. C., Holm, M. B., Beach, S., Schulz, R., & Starz, T. W. (2001). Task independence, safety, and adequacy among nondisabled and osteoarthritis-disabled older women. *Arthritis care & research, 45*, 410-418.

Rogers, J. C., Holm, M. B., & Stone, R. G. (1997). Evaluation of daily living tasks: The home care advantage. *American Journal of Occupational Therapy, 51*, 410-422.

Roth, E. J., Heinemann, A. W., Lovell, L. L., Harvey, R. L., McGuire, J. R., & Diaz, S. (1998). Impairment and disability: Their relation during stroke rehabilitation. *Archives of Physical and Medical Rehabilitation, 79*, 329-335.

Rubenstein, L. V., Calkins, D. R., Greenfield, S., Jette, A. M., Meenan, R. F., Nevins, M. A., et al. (1988). Health status assessment for elderly patients: Report of the society of general internal medicine task force on health assessment. *JAGS, 37*, 562-569.

Scherbaum, Jr., C. A. (2003). *Detecting intentional response distortion on measures of the five-factor model of personality: An application of differential person functioning.* Unpublished doctoral dissertation, Ohio University, Ann Arbor, MI.

Shah, S., Vanclay, F., & Cooper, B. (1989). Improving the sensitivity of the Barthel Index for stroke rehabilitation. *Journal of Clinical Epidemiology, 42*, 703-709.

Shinar, D., Gross, C. R., Bronstein, K. S., & Licata-Gehr, E. E. (1987). Reliability of the activities of daily living scale and its use in telephone interview. *Archives of Physical and Medical Rehabilitation, 67*, 723-728.

Shiotsuka, W., Burton, G. U., Pedretti, L. W., & Llorens, L. A. (1992). An examination of performance scores on activities of daily living between elders with right and left cerebrovascular accident. *Physical and Occupational Therapy in Geriatrics, 10*(4), 47-57.

Silverstein, B., Fisher, W. P., Kilgort, K. M., Harley, J. P., & Harvey, R. F. (1992). Applying psychometric criteria to functional assessment in medical rehabilitation: II. Defining interval measures. *Archives of Physical and Medical Rehabilitation, 73*, 507-518.

Sinoff, G., & Ore, L. (1997). The Barthel activities of daily living index: self-reporting versus actual performance in the old-old (>= 75 years). *Journal of American Geriatric Society, 45*, 832-836.

Skidmore, E. R. (2003). *Functional assessment after stroke: Examining outcome models using the International Classification of Functioning, Disability and Health.* Unpublished doctoral dissertation, University of Pittsburgh, Pittsburgh, PA.

Smith, R. M., & Taylor, P. A. (2004). Equating rehabilitation outcome scales: Developing common metrics. *Journal of Applied Measurement, 5*, 229-242.

SPSS Inc. (2005). *SPSS for Windows (version 14.0).* Chicago: SPSS Inc.

Stone, S. P., Halligan, P. W., & Greewood, R. J. (1993). The incidence of neglect phenomena and related disorders in patients with an acute right or left hemisphere stroke. *Age and Ageing, 22*(1), 46-52.

Tennant, A., & Conaghan, P. G. (2007). The Rasch measurement model in rheumatology: What is it and why use it? When should it be applied, and what should one look for in a Rasch paper? *Arthritis & Rheumatism, 57*, 1358-1362.

Tesio, L. (2003). Measuring behaviours and perceptions: Rasch analysis as a tool for rehabilitation research. *Journal of Rehabilitation Medicine, 35*, 105-115.

Tilley, B. C., Marler, J. R., Geller, N. L., & National Institute of Neurological Disorders and Stroke [NINDS] rt-PA Stroke Trial Study. (1996). Use of a Global Test for Multiple Outcomes in Stroke Trials With Application to the National Institute of Neurological Disorders and Stroke t-PA Stroke Trial. *Stroke, 27*, 2136-2142.

Tsiji, T., Sonoda, S., Domen, K., Saitoch, E., Liu, M., & Chino, N. (1995). ADL structure for stroke patients in Japan based on the Functional Independence Measure. *American Journal of Physical Medicine & Rehabilitation, 74*, 432-438.

Tyson, S., & Turner, G. (2000). Discharge and follow-up for people with stroke: What happens and why. *Clinical Rehabilitation, 76*, 1152-1155.

van Boxel, Y. J. J. M., Roest, F. H. J., Bergen, M. P., & Stam, H. J. (1995). Dimensionality and hierarchical structure of disability measurement. *Archives of Physical and Medical Rehabilitation, 76*, 1152-1155.

van Hartingsveld, F., Lucas, C., Kwakkel, G., & Lindeboom, R. (2006). Improved interpretation of stroke trial results using empirical Barthel item weights. *Stroke, 37*, 162-166.

van Swieten, J. C., Koudstaal, P. J., Visser, M. C., Schouten, H. J. A., & van Gijn, J. (1988). Interobserver agreement for the assessment of handicap in stroke patients. *Stroke, 19*, 604-607.

Wade, D. T., & Collin, C. (1988). The Barthel Index: A standard measure of physical disability. *International Disability Studies, 10*(2), 64-67.

Wolfe, C. D. A., Taub, N. A., Woodrow, E. J., & Burney, P. G. J. (1991). Assessment of scales of disability and handicap for stroke patients. *Stroke, 22*, 1242-1244.

Wolfe, F., Hawley, D. J., Doldenberg, D. L., Russell, I. J., Buskila, D., & Neumann, L. (2000). The assessment of functional impairment in fibromyalgia (FM): Rasch analyses of 5 functional scale and the development of the FM health assessment questionnaire. *Journal of Rheumatology, 27*, 1989-1999.

Woodson, A. M. (2002). Stroke. In C. A. Trombly & M. V. Radomski (Eds.), *Occupational Therapy for Physical Dysfunction* (5th ed., pp. 817-853). Baltimore: Lippincott Williams & Wilkins.

Wright, B. D. (1977). Misunderstaning the rasch model. *Journal of Educational Measurement, 14*, 219-225.

Wright, B. D. (1996). Reliability and separation. *Rasch Measurement Transactions, 9*, 472.

Wright, B. D., & Linacre, J. M. (1989). Observations are always ordinal; Measurement, however, must be interval. *Archives of Physical and Medical Rehabilitation, 70*, 857-860.

Wright, B. D., & Masters, G. N. (1996). Number of person on item strata. *Rasch Measurement Transactions, 16*, 888.

Yavuzer, G., Küçükdeveci, A., Arasil, T., & Elhan, A. (2001). Rehabilitation of stroke patients: Clinical profile and functional outcome. *American Journal of Physical Medicine & Rehabilitation, 80*, 250-255.

Yu, C. H., & Osborn Popp, S. E. (2005). Test equating by common items and common subjects: Concepts and applications. *Practical Assessment Research & Evaluation, 10*(4), 1-19.

Zinn, S., Dudley, T., Bosworth, H. B., Hoening, H. M., Duncan, P. W., & Horner, R. D. (2004). The effect of poststroke cognitive impairment on rehabilitation process and functional outcome. *Archives of Physical and Medical Rehabilitation, 85*, 1084-1090.