Preliminary Development of the Hand and Arm Function Measure for People with Neurological Conditions

By

Namrata Grampurohit

A dissertation submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

University of Washington

2016

Reading Committee:

Deborah Kartin, PhD, Chair

Brian Dudgeon, PhD

Catherine Taylor, PhD

Valerie Kelly, PhD

Kathryn Yorkston, PhD

Program Authorized to offer degree:

Rehabilitation Science

University of Washington

Abstract

Preliminary Development of the Hand and Arm Function Measure for People with Neurological Conditions

Namrata Grampurohit

Chair of the Supervisory Committee:

Deborah Kartin, PT, PhD

Department of Rehabilitation Medicine

Background: Many people with neurological conditions experience difficulty with their hands and arms. Measures for assessing hand and arm function with sound psychometric properties are needed. Understanding patient/client's perceptions of their hand use and the quality and speed of task performance, while critical, are not combined in a comprehensive measure.

Purpose: To develop Hand and Arm Function Measure (HAFM) with stakeholder engagement, that is performance-based, and includes self-report.

Method: A systematic review of literature of hand and arm function measures was conducted and psychometric properties summarized. Four experts, two occupational therapists, a physical therapist and a psychometrician, participated in a focus group discussing the construct, early item bank development, and later participated in item writing and modification. People with stroke (n=7), traumatic brain injury (n=2), multiple sclerosis (n=5) and Parkinson disease (n=6) participated in seven focus groups and eight cognitive interviews, sharing their experiences of hand and arm function difficulties in daily life and during the HAFM administration.

Results: Twenty-two measures of hand and arm function were reported in literature in the past decade. Analysis revealed, they lacked psychometric rigor, ignored the sensory capability of the hand, and were inadequately studied in people with neurological conditions. The construct definition was developed based on the frameworks of International Classification of Functioning, Disability and Health and the principles of Evidence Centered Design. The experts helped define the construct, develop the preliminary item bank, and modify items. The focus groups of people with neurological conditions helped identify items and aspects of hand function that were important to them. All items were pooled and a preliminary set of self-report and performance-based items was developed for use in the cognitive interviews. Based on the cognitive interviews, HAFM items needing modification or suggested for addition/deletion to the HAFM were nominated.

Conclusions: Data gathered from literature review, theoretical and measurement frameworks, and the stakeholders contributed to the face and content-related evidence for validity of the HAFM scores. Plans for future HAFM development include piloting the preliminary item set with people with neurological conditions and ultimately proceeding to item tryouts and large-scale administration.

Plain Language Summary

People with neurological conditions often have problems with their hands and arms. It is important to have tests that measure changes in hand and arm use. New tests need to be developed with care to make sure they give us useful information. Tests should include tasks that people can do and a way for people to tell us about their own experiences with hand and arm use. Right now no tests address both these areas. This is why a new test, the Hand and Arm Function Measure (HAFM), is being developed.

First we read a lot about existing hand and arm function tests. We also got feedback from people with neurological conditions and experts. Four experts had a discussion about hand and arm function and came up with some questions and tasks that could be used in the test. Twenty people with neurological conditions also had group discussions about hand and arm use and eight of them were individually interviewed. They told us about their problems with their hands and arms in daily life and their experiences when they were tested with an early version of the HAFM. They suggested items to include in the HAFM.

We found twenty-two tests of hand and arm function over the past ten years. Those tests were not rigorously developed. They had not been studied in people with neurological problems. Based on theories and feedback from the experts and people with neurological problems, we came up with our own definition or construct of hand and arm function. Items from experts and people with neurological conditions were pooled and a set of items was developed for the HAFM. More input from the experts and people with neurological conditions was used to modify the HAFM items. More work needs be done

on the HAFM including testing more people with neurological conditions and healthy adults.

Human Subject's Application Number: **47121**

Approval Date: **04/24/2014**

Approved Renewal Date: **04/24/2015**

Institutional Review Board: **University of Washington**

Acknowledgements

It is a great pleasure to have the opportunity to thank the many wonderful people in my life whose strong support and dedication, has made this dissertation possible.

First, I would like to express my sincere and deepest gratitude to my advisor and mentor, Dr. Deborah Kartin, who has been a powerhouse of encouragement, guidance and knowledge and mentored my journey to becoming a scientist. Among the many intense phases of the project, I am indebted to Dr. Kartin for the long hours she has spent with supervising and helping with the qualitative data from collection to coding and analysis. This behemoth task would not have been complete without her.

I would like to sincerely thank my mentor, Dr. Brian Dudgeon, who guided my PhD journey in the beginning years, patiently listened to every new idea I presented to him, and continued to guide this dissertation research remotely. Dr. Dudgeon's expertise in occupational therapy and qualitative research provided invaluable guidance for the project.

I feel fortunate to have trained with Dr. Catherine Taylor, in instrument development and am deeply thankful for her guidance throughout my dissertation. I have greatly benefitted from her willingness to share her wisdom and expertise in measurement science and aspire to someday unravel the mysteries of the 'theory of validity' to my students as tactfully as they were unraveled to me by Dr. Taylor. I sincerely appreciate her commitment to teaching and staying on my committee even with the many changes presented through the years.

I am immensely thankful to Dr. Valerie Kelly for sharing her expertise in movement science and neural control and supporting me through this dissertation. I would also like to thank Dr. Kathryn Yorkston for her guidance in organizing the systematic review and qualitative methods. I thank Dr. Nancy Woods and Dr. Basia Belza for their support. I have also been fortunate to be able to consult with Dr. Dagmar Amtmann, Dr. Janet Powell, Dr. Christine Chen, and Dr. Wendy Coster (Boston Rehabilitation Outcomes Center Consultation Services) on this project and would like to thank them for their willingness to share their knowledge with me and for the insightful discussions.

I would like to thank Walter C. and Anita C. Stolov Research Fund for funding this research project. I would like to thank Dr. Peter Esselman, Chair of the Department of Rehabilitation Medicine, University of Washington and all the staff and students that I have worked with over the years for their support, and teaching and research opportunities. I would also like to acknowledge the inspiration I have found in Dr. Jean Deitz's pioneering work in Occupational Therapy at the University of Washington and thank her for her willingness to guide me through the decision to pursue this degree. I would like to thank Dr. Sarah Westcott-McCoy, Dr. Chet Moritz and Dr. Mark Guthrie for support within the Physical Therapy Division. I would also like to thank Victoria Corrin, Laura Robinson, and Mary Garcia for their support.

I would like to acknowledge and thank from the bottom of my heart, Dr. Sujata Pradhan, who said, "You can do it" from my very first day in the PhD program. Her informal mentoring has kept me afloat through the highs and lows, which were in plenty during the last few years. On the same note, I would like to thank the many PhD students and post-docs in the cohorts senior to me for the valuable pieces of advice. I would like to give a shout out to my PhD cohort, Dr. LinYa Hsu, Dr. Sara Morgan, and Dr. Torey Gilbertson for being fabulous friends.

There are many wonderful people in my life outside of the University of Washington who enabled me to complete this dissertation. Foremost, I would like to thank Renee King, for adjusting many different schedules in the hospital to make my part-time clinical work at Evergreen Health work for me through the PhD years. I was fortunate to add Evergreen Health as another clinical site for data collection through the support of Dr. Ted Brown, Dr. Pinky Agarwal, Dr. Marlan Kay, Dr. James Song, Kay Ballen, Ann Zylstra, Kirk Hanson, Susan Watters, Jesika Shah and Sumedha Kukreja.

I would also like to thank my participants in this research project and my patients through the years who have inspired me to search for better tests and outcomes in rehabilitation, making the research pursuits worthwhile. I would like to thank the many therapists who have shared their clinical expertise with me.

Finally, a special thank you to my loving family in US and India, who have always believed in me. I deeply thank my amazing parents, Dr. Dattatraya Grampurohit and Dr. Nirmala Grampurohit, whose own PhD endeavors and successful careers have predisposed me to seek the completion of this degree. I would like to greatly thank my sister, Shweta Grampurohit for her support and volunteering to do the illustrations for the Hand and Arm Function Measure designed in this research. Thanks to my brother-in-law, Dr. Ajinkya More for assisting with the illustrations and for his encouragement.

A special thank you to my enthusiastic sister-in-law Gauri Khandekar and brother-in-law Dr. Sidharth Chiravarambath for their ardent support and patience, letting me work on my dissertation writing through the Thanksgiving holidays. I am thankful for the smiles shared by Om and Meera to liven up some of the toughest days. Thanks to my parents-in-law Saktan Chiravarambath and Hemlata Saktan for celebrating my achievements.

I would like to thank my son, Neil Chiravarambath for his love, saying ‘its okay’ when things were not perfect, his interest in always looking at my new drafts, and keeping me excited about graduating from school. I would like to thank my daughter Keya Chiravarambath for her love, the many hugs that gave me strength and her help in organizing my dissertation papers and books.

I am eternally grateful to my beloved husband, Kapil Chiravarambath, who is the reason I was able to start and finish this degree. He has nurtured my aspirations, given me strength and helped raise a beautiful family with twenty years of love and togetherness.

For Kapil and my family

Table of Contents

List of Figures

List of Tables

List of Appendices

Chapter 1: Introduction

Skillful hand use is essential to daily tasks, self-management, community activities, leisure, and employment. We use our hands to grasp, hold, manipulate objects and tools, and to support our bodies (Radomski & Trombly, 2007). We also use hands for uniquely human functions from identifying objects and gestures to sign language and reading Braille. Functions of the human hand are many and varied. The human hand function ranges from sensory to motor functions (Jones & Lederman, 2006). It can be viewed as a continuum from tactile sensing (e.g., passive touch) to active haptic mode (e.g., perception and manipulation of objects using the senses of touch and proprioception) to prehension (e.g., tool use) to non-prehensile skilled movements (e.g., keyboarding) (Jones & Lederman, 2006). The neural control of the hand is far more complex involving grasp and reach specialized regions in the brain along with multiple interconnections to enable cognitive, visual and emotional control of hand movements.

We need hands and we take them for granted until impairments make us aware of their true potential (Napier & Tuttle, 1993). Hand function impairments cause difficulty participating in life roles (Al Snih, Markides, Ottenbacher, & Raji, 2004; Radomski & Trombly, 2007). Neurological disorders such as stroke, traumatic brain injury (TBI), Parkinson disease (PD), and multiple sclerosis (MS) cause hand function impairments (Kandel, Schwartz, Jessell, Siegelbaum, & Hudspeth, 2012). These disorders are commonly encountered by therapists in the clinic and are the focus of this research study. Stroke and TBI are non-progressive conditions and MS and PD are progressive conditions affecting hand function. The differences in the way these disorders can affect arm and hand function is presented in Table 1.1.

After stroke, on discharge from the hospital, often arm-hand function is not recovered (Tyson & Turner, 1999, 2000). Four years after stroke, 67% of the patients experience the non-use or disuse of the affected arm as a major problem (Broeks, Lankhorst, Rumping, & Prevo, 1999). In severe TBI, hand function deficits related to motor impairments are commonly encountered (NIH Consensus Development Panel on Rehabilitation of Persons With Traumatic Brain Injury, 1999). For people with MS, hand function impairments may be quite disabling (Feys, Romberg, Ruutiainen, & Ketelaer, 2004) and affect their quality of life (Yozbatiran, Der-Yeghiaian, & Cramer, 2008). PD and MS both can be treated with disease modifying drugs that can slow down the progression of the disease, thereby reducing the impact on hand function in daily activities.

Therapists and other professionals need hand and arm function measurement tools that are relevant to individuals with neurological condition (Kraft et al., 2014; Velozo, Seel, Magasi, Heinemann, & Romero, 2012). The term 'measure' is defined as 'a system or scale of standard measuring units' when used in reference to hand and arm function measurement. Clinicians need reliable and valid measures of hand function so they can monitor progress, set goals, determine effectiveness of intervention, and seek reimbursement for therapy services (Carmeli, Patish, & Coleman, 2003; Lemmens, Timmermans, Janssen-Potten, Smeets, & Seelen, 2012). In order for a measure of hand function to be most useful for these purposes, test items need to be well developed and the scores need to be reliable and valid for the stated purpose (Lehman et al., 2011). In the words of Elaine Ewing Fess – an expert in hand therapy, author of Functional Tests in the book *Rehabilitation of Hand and Upper Extremity* (Skirven, Osterman, Fedorczyk, & Amadio, 2011) – on assessment of hand,

> *"A thorough and unbiased assessment procedure furnishes information that helps predict rehabilitation potential, provides data with which subsequent measurements may be*

compared, and allows medical specialists to plan and evaluate treatment programs and techniques." (p. 152)

The existing hand function measures can be categorized into performance-based measures and self-reported measures. Some performance-based measures such as the Jebsen Taylor Hand Function Test (Stern, 1992) and the Nine Hole Peg Test (Oxford Grice et al., 2003) can be used for multiple diagnostic populations (Asher, 2007). Other performance-based measures are diagnosis specific such as the Wolf Motor Function Test (D. Morris, Uswatte, Crago, Cook, & Taub, 2001) or Action Research Arm Test (Yozbatiran et al., 2008), which are used to assess hand impairments in individuals post-stroke (Velstra, Ballert, & Cieza, 2011). More recently, self-report measures for assessing hand function have been developed such as the Motor Activity Log for survivors of stroke (Uswatte, Taub, Morris, Vignolo, & McCulloch, 2005). Measures developed for people with stroke cannot be used in other diagnostic conditions unless there is evidence to support their use. Moreover, these currently available measures have many limitations such as small normative sample sizes, poor evidence for validity of scores at high and low end of impairments, lack of items that assess activity and participation, and lack of psychometric studies in the population of interest (Ashford, Slade, Malaprade, & Turner-Stokes, 2008). Moreover, the performance measures require significant time for administration and may have costs associated with purchase. Thus, there is a need for the conceptualization and development of a measure of hand and arm function to overcome these limitations.

Recent research points towards the need for two components to assess hand function in the clinic: 1) perceived hand use and 2) the quality and speed of task performance (Rallon & Chen, 2008). Currently no measure exists that has fully addressed both a performance component and a self-report component in a single measure (Asher, 2007). This dissertation begins the process of development of a new measure of hand and arm function, i.e., Hand and Arm Function Measure (HAFM) that includes both these components. This new measure of hand and arm function will have the potential to shed light on the importance of using self-report along with task performance. It is intended be practical, as there are a multitude of daily living tasks that can be measured via self-report but cannot be administered and tested reliably within the evaluation time frame of most clinical settings.

There are many stages in the development of a new measure. Developing an item bank is recognized as one of the critical steps in this process (D. Cook & Beckman, 2006; K. Cook et al., 2007). Lehman et al. (2011) have conducted and documented a similar study in which they developed an upper extremity item bank for patient reported outcome measure in musculoskeletal conditions with two uni-dimensional constructs of gross upper extremity and fine hand use.

Grounding of a new measure in theory and evidence requires developing the construct or idea behind its development (Messick, 1989). Researchers in education and psychology have developed a systematic approach to guide this process called Evidence Centered Design (ECD) Framework (Mislevy, 2011; Mislevy, Almond, & Lukas, 2003). This framework provided structure to the process of developing the HAFM. This framework suggests literature review, expert input and participant information to be utilized in the process of development of a measure before piloting the items. There are only a few hand function measures that have been developed in a systematic manner (O'Dell et al, 2013) and none that have utilized state-of-the-art measurement approaches such as ECD Framework.

The ECD framework emerged from an effort (Mislevy, 2007; Mislevy et al., 2003) to bridge the gap between the conceptualization of construct-related evidence for validity and practical ways in which researchers can implement the idea of supportive arguments for the interpretation and use of test scores. ECD is based on the more recent work by Samuel Messick (Messick, 1994). ECD has been used in diverse assessments such as teacher certification examinations and commercial vehicle driver assessments (Mislevy, 2011). The assessment argument is structured in layers of ECD, which are organized serially however the work continues with iterations back and forth between theory, construct development, assessment development, and evidence. The layers of ECD are as follows:

1. Domain analysis: This is the first step in assessment design where analysis involves gathering information about the domain in the context of assessment. Here, domain implies the construct or abstract idea. The analysis also goes deeper into the content, concepts, and terminology and representational forms that people working in this domain use. The knowledge representations important for the domain include knowledge, skills and attitudes that are relevant. A common way of domain analysis is task analysis, which has been practiced widely in rehabilitation science.
2. Domain modeling: This involves the conceptualization process. In domain modeling the information and relationships discovered in domain analysis are organized in a narrative form. Some examples of forms of representation of domain modeling are claims and evidence worksheets showing the relationship between what claims are to be made based on assessment scores and what evidence is gathered to support the claim. Toulmin diagrams and design patterns are other ways of domain modeling that help to structure the evidence (Mislevy et al., 2003). This step facilitates the link between design patterns, Toulmin arguments and the upcoming step of a Conceptual Assessment Framework.

3. Conceptual Assessment Framework: This step involves writing the test and task templates, test and task specifications and scoring algorithms. Messick (Messick, 1994) suggested organizing this step in three models from the perspective of student (patient or client), evidence and task.
4. Assessment implementation: The making of the actual test that includes authoring tasks, scoring details, implementing statistical models, piloting and finalizing evaluation procedures and producing test materials and presentation environments (paper pencil versus computer based) that will be utilized.
5. Assessment Delivery: Mislevy (2007) has laid out four-process delivery architecture. This assessment delivery process is where 1) the therapist selects the task (activity selection), 2) patient/client interacts with tasks (presentation process), 3) performances are evaluated (evidence identification process), 4) feedback is created and reports are generated (evidence accumulation process).

These steps are guidelines and there is a planned redundancy in this design. For example, some decisions involved in domain modeling and conceptual assessment framework may overlap. However, this further strengthens the process and assures that the test designer carefully examines the background and nature of the construct before delivering the final product. The ECD framework also structures the terminology used for the measurement concepts. The older terms ‘reliability and validity of the measure’ are now appropriately modified to ‘evidence for reliability and validity of scores of the measure.’ This is a major change in the field of measurement brought about by the focus on evidence. Thus, ECD provides structure to the research process of developing a new measure in many different ways.

Some other assessment frameworks have been developed in recent years such as the patient reported outcome measure development framework given by the Federal Drug Administration (FDA) (U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research, U.S. Department of Health and Human Services FDA Center for Biologics Evaluation and Research, & U.S. Department of Health and Human Services FDA Center for Devices and Radiological Health, 2006). The steps in the FDA framework are presented in a circular flow diagram and include: 1) identify concepts and develop conceptual framework, 2) create instrument, 3) assess measurement properties, and 4) modify instrument. However, the ECD framework was selected over the FDA framework as it allows for more conceptual and domain specific development of the measure at various stages from construct to delivery; it ties all of the steps together in the diagrams that support the construct; and there is no intention to seek FDA approval. That said, the basic methods of this research align with the ECD and FDA frameworks.

The overall purpose of this research was to ultimately develop a psychometrically robust hand and arm function measure for people with neurological disorders that assesses both performance-based and self-report components. The performance-based component assesses tasks and self-report component assesses daily living items targeted towards activity and participation. This dissertation research begins the process of developing this measure for clinical use by rehabilitation professionals. Developing a new measure for arm and hand function is a lengthy process that needs a systematic approach. It is the thesis of this work that to develop a new measure of arm and hand function, qualitative information from stakeholders such as people with neurological condition, clinicians who are involved in care, and measurement experts is needed to inform item bank development.

The objectives of this study were:

- To review the literature regarding hand and arm function and existing hand and arm function measures to inform the construct definition.
- To develop the construct definition for a new arm and hand function measure.
- To synthesize qualitative information from experts for item bank development, item writing and item modification.
- To synthesize qualitative information from people with neurological condition to develop item bank and during HAFM administration.

The chapters are organized in the order in which the research studies were conducted. Chapter 2 titled 'Literature Review' provides a background of literature related to classification of hand function, neural control of arm and hand, and a systematic review of hand function measures. Chapter 3 titled 'Construct definition and Expert Panel Study' explains the construct of arm and hand function and explains the expert panel study of the construct and item development. Chapter 4 titled 'Focus groups and Cognitive Interviews Study' explains the qualitative study involving people with neurological conditions and the process of item bank development. Chapter 5 titled 'Discussion and Future Research' presents an overarching discussion of the research studies conducted to-date, explains the process adopted for piloting of HAFM, and presents the long-term research plan for developing a psychometrically robust HAFM.

The term 'hand and arm function' in this dissertation was used generally to mean the function of the entire upper extremity and includes shoulder, arm, elbow, forearm, wrist, and finger movements for function. The terms 'hand and arm function,' 'hand function,' 'upper extremity function,' and 'upper limb function,' are used interchangeably in the literature and in this dissertation. In this dissertation, although every attempt was made to use the term 'measure' to refer to the assessments, some other terms such as 'test', 'questionnaire,' 'index,' 'scale,' 'tool,' 'inventory,' 'instrument,' or 'assessment' may be used interchangeably. In this dissertation, 'tool' and 'instrument' were terms used not only for mechanical measuring devices such as dynamometer, but also occasionally used for questionnaires. The use of the term 'outcome' meant the change due to the effect of a therapy, training, or due to other effects. The terminology used for describing measurement properties is defined in Appendix A.

To summarize, the development of a new measure of hand and arm function should be grounded in theory and experiences of the stakeholders to overcome the limitations of the currently available measures. The ECD framework utilized for scale development helps systematically address each area of development to achieve the ultimate purpose of developing a psychometrically robust measure of hand and arm function.

Table 1.1. Descriptions of Arm and Hand Issues in Neurological Conditions

Neurological condition	*Progression*	*Arm and hand issues*
Stroke	Non-progressive	Mostly unilateral deficits with motor and sensory involvement, learned non-use, low tone initially, followed by spasticity.
Traumatic Brain Injury	Non-progressive	Unilateral or bilateral deficits, associated vision and coordination deficits, most recover fully in 6 to 12 months, and/or low, high or fluctuating tone.
Parkinson Disease	Progressive	Tremors on both sides early on, incoordination, slow movements, involuntary movements, rigidity and/or weakness.
Multiple Sclerosis	Progressive	Mostly bilateral sensorimotor impairments along with spasticity, fatigue and/or vision impairments.

Chapter 2: Literature Review

Abstract

Background: The human hand function has been classified in different ways to identify the unique abilities of the hand. The neural control of hand is complex and the contemporary theories on voluntary control of the hand, guide the development of the measure designed for people with neurological conditions. When developing a new measure of hand function, it is important to have an understanding of human hand function, the neural control of the hand and what hand function measures currently exist. There are many measures of hand function developed over the years. Only a few have gone through the rigor of psychometric testing and many lack evidence for validation in the neurological conditions of stroke, multiple sclerosis, traumatic brain injury and Parkinson disease.

Objective and Method*:* An in-depth literature review of hand function classification and neural control theories was conducted with the objective of identifying the models that would provide the theoretical foundation for the Hand and Arm Function Measure (HAFM). A systematic review of literature was conducted with the objective of identifying and describing the current hand function measures for people with neurological conditions. The secondary objective of the literature review was to map these hand and arm function measures to the aspects of hand function addressed by the measures.

Results: The literature review identified Jones and Lederman's classification as a particularly useful model for classifying human hand function as it takes into account the sensory capability of the hand. The International Classification of Function, Disability and Health (ICF) also influenced the structure of the HAFM. Both these theoretical foundations for the measure are discussed in detail in this review. The influence of the contemporary models of neural control is also discussed in depth for their role in shaping the development of the measure. The

systematic review identified 22 hand function measures used in 86 studies in the past decade (January 2005 to July 2014). None of the measures had a combination of self-report and performance-based sections. The most commonly used performance tests were the Action Research Arm Test, the Fugl-Meyer Assessment, and the Wolf Motor Function Test (WMFT). The most commonly used self-report measures were the Motor Activity Log and the Stroke Impact Scale –Hand domain. Most of the measures lacked psychometric rigor, did not include sensory capability of the hand, and were not adequately studied in neurological conditions.

Conclusions and Significance*:* The new measure is grounded in the theoretical foundations of the ICF, models of classification of hand function, and contemporary neural control theories. The systematic review on hand function measures highlights the need for a new measure of hand function that is psychometrically robust, provides evidence for validation in people with neurological conditions, developed systematically, and contains self-reported and performance-based components integrated in the measure.

Introduction

The theoretical foundation for measuring upper extremity function in neurological disorders is based on what we know about human hand function, the neural control of the arm and hand, and the measures developed in the past. A critical appraisal of existing literature on hand function tests and their psychometrics provides for an analysis of contemporary views around hand function measures. This literature review chapter is organized into: models of classification of hand function, a review of the neural control of the arm and hand including brain areas controlling movements, and a structured review of hand function measures.

Models of classification of hand and arm function

Human hand function is complex and can be classified in numerous ways. In occupational therapy (OT) literature, hand function has been broadly classified into gross motor (i.e., movements involving large muscles of the arm and hand) and fine motor function (i.e., movements involving small muscles of the hand) (Radomski & Trombly, 2007). Also, prehension that involves finger movements and grasp that involves the use of the palm have been considered separate aspects of hand function (Pendleton & Schultz-Krohn, 2006). The diverse range of postures that the hand can assume, have been described by Jones and Lederman (Jones & Lederman, 2006). These postures include 16 different power and precision grasps from the original taxonomy presented by Cutkosky & Wright for robotic hands (Cutkosky & Wright, 1986; Venkataraman & Iberall, 2012). Hand function has also been classified into reach, grasp, and manipulation by motor control theorists (Shumway-Cook & Woollacott, 2011).

In the late 1970s, Sollerman and Sperling (1978) developed a handgrip classification system based on 4 finger grips (i.e., pulp pinch, lateral pinch, tripod pinch, and five-finger pinch) and 4 volar grips (i.e., diagonal volar grip, transverse volar grip, spherical volar grip, extension grip). Napier and Tuttle (1993) broadly classified hand function into prehensile and non-prehensile functions. Later, Jones and Lederman (2006) presented the sensorimotor continuum with the five categories that included tactile sensing, active haptic mode, prehension, non-prehensile skilled movements and transport. In motor control literature (Shumway-Cook & Woollacott, 2011) the grasping patterns are reported to be a function of location, size and shape of the object however, in contrast Napier and Tuttle (1993) argue that the intended activity is the most important consideration. Jones and Lederman's classification that includes sensory aspects of the hand and the transport function of the arm was adopted for examination of hand function measures in the systematic review presented later in this chapter.

In Jones and Lederman's (2006) classification, tactile sensing refers to the passive detection of superficial sensation applied to the skin. Active haptic mode refers to the active movements involved in detection of properties of objects using proprioception. Prehension includes hand functions that involve various grips (e.g., power, precision, hook) and pinch along with the specialized function of opposition. Non-prehensile skilled movements include pushing, pointing, turning and carrying motions. Transport includes proximal joint functions needed for reach to mouth, overhead, forward, and back. Pediatric occupational therapy literature has further expanded the function of pinch to include in-hand manipulation (Humphry, Jewell, & Rosenberger, 1995). In-hand manipulation is the ability to hold and move an object within one hand and is divided into three major categories: rotation, translation and shift. All Jones and Lederman's categories may include 'stabilization' which involves the ability to hold the object using the ring finger and little finger while moving an object with the thumb, index and middle fingers. Translation is the ability to move objects from the fingertips to the palm or palm to the fingertips such as moving coins to place in vending machine. Shift is the ability to move an object in a linear manner with the fingertips such as fanning cards in the hand. Rotation is categorized as simple and complex. Simple rotation is the ability to turn an object around the pads of the fingers and thumb. Complex rotation is the ability to turn an object end to end such as flipping a pencil from writing end to eraser end.

Many of these aspects of hand function are impaired in people with stroke, multiple sclerosis (MS), traumatic brain injury (TBI), and Parkinson disease (PD). In order to better understand and measure hand function in people with these conditions, it is helpful to have an understanding of the theoretical foundation of the neural control of arm and hand.

Neural Control of Hand and Arm

The neural control of arm and hand is not fully understood. The voluntary control of the arm and hand is different from reflexive and patterned movements of the nervous system. Models have been developed by contemporary theorists to explain mechanisms that can form the theoretical basis of the new measure. Some measures like the Fugl-Meyer Assessment were developed based on the Twitchell and Brunnstrom's concept of sequential stages of motor return in the hemiplegic stroke patient (Brunnstrom, 1966; Gladstone, Danells, & Black, 2002) and assesses synergistic movements. The Motor Assessment Scale (MAS) was developed based on Carr and Shepherd's (Carr, Shepherd, Nordholm, & Lynne, 1985) Motor Relearning Principles.

The new measure developed in this research study, i.e., Hand and Arm Function Measure (HAFM), is based on the current concepts in neural control of individual-task-environment model (Shumway-Cook & Woollacott, 2011) and Shadmehr and Krakauer's (2008) model from computational neuroscience. Shumway-Cook and Woollacott (2011) describe the influence of the individual, task and environment and their interaction with each other in feedback and feed-forward loops. According to this model, an individual and their central nervous system, act in the perception-cognition-action series to solve the problem of picking up a glass of water. The peripheral receptors in hand are involved in sensing and the primary and secondary sensory cortical areas of the brain are involved in perceiving. The parietal association areas in the brain are involved in interpreting and the pre-frontal cortex and association areas are involved in conceptualization. The supplementary motor cortex, the basal ganglia and the cerebellum are involved in developing the movement strategy and the primary motor cortex activates the action sequence. The motor neurons, muscles and joints then execute the action. Shadmehr and Krakauer's (2008) model states that an individual performs movements to achieve a rewarding state. This model accounts for the role of the subcortical structures such as basal ganglia in making an important contribution to movement. This model proposes that the individual will accomplish a given task in the most efficient strategy, path or trajectory with the least cost and most reward. Applying the concepts of this model, a predetermined path of movement for a task on a test item may not be strictly followed by the patient/client if the cost and rewards change.

These models bring the focus of the HAFM on the individual with self-reported items, the task with complexity and time elements, the environment with standardized and familiar equipment, the motor planning aspects of movement with qualitative assessment by the therapist, and the sensory aspects with active assessment combined with movement. Other than these two models, the recent research into reach and grasp regions (Buneo, Jarvis, Batista, & Andersen, 2002; Hwang, Hauschild, Wilke, & Andersen, 2012) in the brain, the two areas that control reach and grasp functions of the arm and hand respectively, also inform the need to test these aspects in the new measure. Thus, prehensile and non-prehensile aspects are included in the measure.

Review of Hand and Arm Function Measures

Assessment of the upper extremity is a vital aspect of clinical evaluation for rehabilitation clinicians working with people with neurological conditions. There are numerous measures of upper extremity function in literature. However, few of them have undergone rigorous psychometric analysis to be used with confidence in research and clinic. Some of the older tests have not stood the test of time and are largely not used in the clinic, e.g., Rivermead Motor Assessment (RMA) (Lincoln & Leadbitter, 1979) . The existing literature reviews on upper extremity function measures conducted in recent years have been primarily related to stroke (Ashford et al., 2008; Ashford & Turner-Stokes, 2013; Lemmens et al., 2012) and do not provide a clear picture of the measures used for other neurological conditions. Other reviews for diagnostic groups such as multiple sclerosis have been conducted but they are narrative reviews that include selected measures (Kraft et al., 2014). No reviews specifically targeting upper extremity measures in traumatic brain injury or Parkinson Disease have been found. Other reviews have focused on classification of the measures into the International Classification of Functioning, Disability and Health (ICF) domains and include orthopedic and neurological conditions (Metcalf, Adams, Burridge, Yule, & Chappell, 2007; Velstra et al., 2011). One review (Velstra et al., 2011) reported frequently used upper extremity outcome measures in stroke and other musculoskeletal conditions, but the selection of outcome measures was arbitrary. Thus, in order to understand the current state of hand function measures in the selected neurological conditions, an in-depth literature search was warranted.

Objective. The objective of this systematic review was to identify and describe the commonly used upper extremity function measures for people with the neurological conditions of stroke, TBI, PD, and MS. A secondary objective was to map these measures to the domains of hand function targeted by the measures.

Method. ***Data sources.*** The PubMed database was searched for studies with the following keywords: ‘hand,’ ‘arm,’ ‘function,’ ‘measure,’ ‘stroke,’ ‘multiple sclerosis,’ ‘Parkinson’s disease,’ and ‘traumatic brain injury’ in the following combination.

> *(hand AND arm) AND function AND Measure AND (stroke OR multiple sclerosis OR Parkinson's disease OR traumatic brain injury) AND ("last 10 years"[PDat] AND Humans[Mesh] AND English[lang])*

Inclusion and exclusion criteria. The inclusion criteria were: measures for adults (19 years of age or older, pre-set in PubMed) with either acute, subacute or chronic stroke, PD, TBI or MS; full text studies in the English language published within the last 10 years (January 2005 to July 2014); and hand function measures with studies of evidence for the reliability and validity of scores. The exclusion criteria were: studies related to pediatrics and non-clinical measures such as robotic assessments and imaging studies with no clinical measures. Instrumented measures that need special measurement devices (e.g., hand held dynamometry) were excluded due to lack of information about activities of daily living. Animal studies were also excluded.

Results. PRISMA guidelines (Preferred Reporting Items for Systematic Reviews and Meta-analysis) are used for reporting of results of this review (Moher, Liberati, Tetzlaff, Altman, & PRISMA Group, 2009). PubMed database search yielded 129 studies. After screening title and abstract, three studies related to pediatrics and 40 studies that did not meet criteria for hand function measures were excluded. Eighty-six studies were selected for full text review and of these 34 hand function measures were identified. Ten measures that were newly developed with limited or no evidence for the reliability and validity of scores and two instrumented measures (i.e., Jamar® dynamometer and Pinchmeter) were then excluded from detailed analysis. This exclusion was based on the initial criteria set for the measures. Figure 2.1 presents the PRISMA flow diagram. A total of 22 measures out of 34 were selected for detailed description in this review. Seven newly developed measures of hand and arm function will be discussed briefly.

Data were extracted from the available studies related to hand function measures in general and psychometrics in particular. Cross references for hand function measures reporting development, evidence for the reliability and validity of scores were used for the purpose of constructing detailed descriptions of the measures in this review. These upper extremity function measures are subsequently described based on their construct, hand function (i.e., prehensile, non-prehensile, transport, tactile sensing and active haptic mode), ICF domain, sample item and response categories, description of population that these measures were developed for and the neurological population for whom they have been used (e.g., stroke, MS, PD and TBI), psychometric properties (e.g., evidence for the reliability, validity and responsiveness of scores), burden of administration, and the number of studies identified in the review that have used the measure.

Through this systematic review 22 hand function measures were identified which were self-report measures and performance-based measures. In the following results, the self-report measures will be reported followed by performance-based measures, along with a brief overview of newly developed measures.

Self-report measures. Seven out of 22 hand function measures found in this review were self-report measures. The details of these measures are outlined in Table 2.1 and 2.2. The most commonly used self-report measures were the MAL (17 studies) and the Stroke Impact Scale – Hand Domain (SIS-HD) (11 studies). The seven self-report measures are presented in the following section from older to newer measures based on the first published study found in the literature.

Disabilities of Arm Shoulder and Hand (DASH). The DASH (Hudak, Amadio, & Bombardier, 1996) is a self-report measure designed to evaluate disorders of the upper limbs and monitor change of function over time. This instrument was developed by the Institute for Work and Health and the American Academy of Orthopedic Surgeons with a systematic two stage process of item generation and item reduction (Hudak et al., 1996). The first stage of development involved input from clinicians and patients and a review of 13 relevant outcome measures in musculoskeletal conditions. The second stage involved item reduction by three content experts, secondary review by a panel of 15 experts for content and face validity evidence and item importance along with pre-testing on 20 individuals with upper extremity musculoskeletal problems (Hudak et al., 1996). Further item reduction was conducted by psychometric analysis. This 38-item scale includes optional eight high-performance items and also has a shorter version (Quick DASH) with 11 items (Mehta, Macdermid, Carlesso, & McPhee, 2010). It can take anywhere from 5 to 30 minutes to complete depending on the type of scale and requires no training. The difficulty with hand function over the past week was evaluated in the measure and patients/clients were asked to rate their best estimate of performance if the activity was not performed over the past week. The wording for response category of each type of item varies with the nature of the item (e.g. symptoms ask about severity or work related tasks ask about difficulty). The measure cannot be scored if more than three missing items exist.

The DASH questionnaire asks about symptoms as well as the ability to perform certain daily activities. DASH includes activities related to sensation (e.g., tingling), transport (e.g., reach), prehension (e.g., heavy and light) and non-prehensile (e.g., pushing and carrying) tasks. The hand function aspects not addressed by this measure are active haptic sensing, in-hand manipulation, and quality and speed of task performance. The DASH assesses difficulty on a 5-point scale with daily activities for unimanual and bimanual tasks irrespective of how the task is done. Strong psychometrics, cultural adaptations and iterations to develop the Quick-DASH have been implemented (Kennedy et al., 2013) for musculoskeletal conditions. The DASH was the only measure that included the work and recreational tasks as separate modules.

The DASH has excellent evidence for reliability of scores (i.e., test-retest, inter-rater and internal consistency) in other upper limb conditions but this has not been established for neurological conditions. Face and content-related evidence for validity of scores was found to be poor specifically in stroke and not established for other neurological conditions (Lannin, McCluskey, Cusick, Ashford, & Ross, 2010). In spite of being a well-developed measure for arm and hand musculoskeletal conditions, the DASH was reported in only three out of the 86 studies found in this systematic review. One of the major reasons for this lack of use in neurological conditions is the absence of research studies to provide evidence of validity of the scores of the measure with patients/clients with neurological conditions (Baker 2015). In addition, the DASH cannot differentiate function on right or left side. People with neurological conditions have unequal distribution of deficits and it is vital to evaluate the ability of each side separately particularly related to symptoms.

ABILHAND. The ABILHAND (Penta, Thonnard, & Tesio, 1998) is a self-report measure of manual ability for adults with upper limb impairments. It measures a person's ability to manage daily activities that require the use of the upper limbs. The ABILHAND has been validated in chronic stroke (age 16-80) and is available in English, French, Dutch, Italian and

Swedish (Penta et al, 1998, 2001). The ABILHAND was reported in two out of 86 studies in this review. This measure was developed in Sweden and has not been used extensively in research. It is an exclusively bimanual measure with 23 daily activities that focus on object handling (i.e., prehensile movements) and takes about 10 minutes to administer (Alt Murphy, Resteghini, Feys, & Lamers, 2015). It has items such as threading a needle that are potentially not appropriate for people with neurological conditions. The non-prehensile and transport functions of the hand are not assessed by this measure. The ABILHAND assesses the difficulty with hand function on a 3-point scale irrespective of the time frame such as past week or past month. The ABILHAND uses modern Item Response Theory (IRT) techniques of Rasch analysis to provide a score if data is entered in the online portal. Paper based version is available for download along with detailed instructions. Each activity is weighted by logits on the Rasch scale for its online version. A logit is a unit of measurement on an interval level of measurement to report relative differences between hand function ability estimates and item difficulties. Evidence for internal consistency is established and order of difficulty of items has been confirmed by Rasch analysis (Penta, 2015; Penta, Tesio, Arnould, Zancan, & Thonnard, 2001; Simone, Rota, Tesio, & Perucca, 2011). No studies were found that reported on intra-rater, inter-rater and test-retest evidence for reliability of scores of the ABILHAND. Concurrent and criterion-related evidence for validity of scores has been established by Penta et al (2001) and Simone et al (2011), along with evidence for responsiveness of the measure. Face validity evidence is not reported.

Motor Activity Log (MAL). The MAL (van der Lee et al., 1999) is a semi-structured interview to assess arm function exclusively for people with stroke and was developed as part of the constraint induced movement therapy research trials. It is a commonly used measure identified in 17 out of 86 final studies in this systematic review. Individuals are asked to rate quality of movement (QOM) (also known as How Well Scale) and the amount of use (AOU) during daily functional tasks. It has unimanual and bimanual tasks related to prehensile, non – prehensile, and gross motor functions related to transport. The original version has 30 tasks and versions with 45, 28 or 14 tasks are also available. The MAL is scored on a 6-point ordinal scale from 0 (weak arm was not used) to 5 (weak arm used as well as before stroke) (Uswatte et al., 2005). It was developed for assessing real world hand use during trials for constraint induced movement therapy and psychometric studies are incomplete as the population studied is restricted to people enrolled for the clinical trials. Although the internal consistency and test-retest evidence for reliability of scores were well established, the minimal detectable change and minimal clinically important difference were not reported, reducing the overall utility of the scores of this measure.

Stroke Impact Scale- Hand Domain (SIS-HD). The SIS (Duncan et al., 1999) was developed as a comprehensive measure for assessment of health status following stroke with hand function domain as one of the eight domains assessed. This is a commonly used measure and was included in 11 out of 86 studies reviewed in this systematic review. Five items in the hand section on this measure include carrying heavy objects (i.e., bag of groceries), turning a doorknob, opening a can or jar, tying a shoe lace, and picking up a dime. There are other items not in the hand domain section of this measure that are related to hand function, for example cutting food with a knife and fork, dressing the upper body, bathing, clipping toenails, and completing light household chores. These items are prehensile movements and non-prehensile (e.g., carry) functions. Transport functions such as reach are not directly addressed in the measure. The items are rated by difficulty on a 5-point scale (1=Could not do at all, 5=Not difficult at all) and are related to performance on a typical day. Psychometric properties of SIS have been studied for the overall measure and for hand function domain specifically (Duncan et al., 2002). The overall scale has excellent evidence for test-retest and inter-rater reliability of scores (Carod-Artal, Ferreira Coral, Stieven Trizotto, & Menezes Moreira, 2009; Duncan et al., 2002). The SIS-HD has adequate concurrent evidence for validity of scores and can discriminate between degrees of stroke severity with adequate correlations with Barthel Index and functional status (Carod-Artal et al., 2009). The measure is sensitive for minor and moderate stroke from 1 to 3 months and 1 to 6 months post-stroke (Duncan et al., 1999). Floor and ceiling effects have also been established (Duncan et al., 1999). In spite of extensive research into psychometric properties, face and content-related evidence for validity of scores for the hand domain have not been satisfactorily reported. The SIS-HD in itself is not a comprehensive measure of hand function for tracking progress due to the lack of hand function items and a scatter of items related to hand in other domains. Further, the SIS-HD was specifically developed for people with stroke (Duncan et al., 1999; van der Lee, Beckerman, Knol, de Vet, & Bouter, 2004) and not studied in other neurological conditions.

Manual Ability Measure (MAM). The MAM (Chen, Granger, Peimer, Moy, & Wald, 2005) is an occupation-based and patient-centered outcome tool that measures manual ability using a self-report format. The MAM was reported in one out of 86 studies in this systematic review. The measure was developed by filtering items from the DASH, TEMPA, ABILHAND and Sollerman Hand Function Test (Chen, 2005). The MAM has 16, 20 or 36-item formats. It includes unimanual and bimanual tasks with primarily prehensile movements related to daily activities. It measures item difficulty on a 4-point scale (1 = I cant do, 4 = Easy). It has a NA (Not applicable) rating that is used when a person never did the task even before his/her hand condition. The Rasch Rating Scale Analysis method was used to develop and refine the instrument. Therapist input was obtained for scale development however patient input is lacking in item development. There is no time frame specified for the patient to reference when answering questions. The MAM (36-item version) was developed for musculoskeletal conditions and later validated for people with neurological conditions including stroke (Chen & Bode, 2010). Although, the construct-related evidence for validity of scores for this measure is studied, it is limited by small sample size overall with two studies found for psychometrics. The evidence for reliability of scores, measurement error, and interpretability of the scores has not been reported.

NeuroQOL-UE. The NeuroQOL (Cella, 2015) is a self-report of health related quality of life. Upper extremity function is one of the 17 domains included in this measure. The NeuroQOL was reported in one out of 86 studies in this systematic review. This is the only measure currently known to be available in Computer Adapted Testing (CAT) format. The measure consists of fine motor and ADL items in the CAT format with variable number of items. Paper-based short forms (8-10 items) are also available across all domains. The items are rated on a 5-point scale based on intensity. Scoring is via computer or can be calculated using tables in the NeuroQOL user manual for each short form. Items have been targeted to people with neurological conditions however the NeuroQOL lacks the sensitivity to change, reducing its utility in the clinic (Kraft et al., 2014). Psychometrics have been established for the NeuroQOL as a generic measure and no studies specifically related to the UE domain were found. With the available psychometrics, it is unclear if using the UE domain by itself will be beneficial for neurological population. Information presented in Table 2.1 and 2.2 is generic data (not specific to upper extremity function sub-section) from NeuroQOL final report (Cella, 2015).

Patient Reported Outcome Measurement Information System® Physical Function Bank for Upper Extremity (PROMIS® – UE). The PROMIS® – UE (Hays et al., 2013), funded by the National Institute of Health (NIH), is a self-reported measure of upper extremity function developed using mixed qualitative and quantitative methods. This measure was reported in 2 out of 86 studies in this review. The upper extremity function short form contains 16 items targeting ADL tasks including unimanual and bimanual activities rated on a 5-point scale of difficulty. Psychometric studies have been done on diverse populations that include people with neurological conditions. Concurrent evidence for validity of scores with DASH ($r = 0.726$) for the PROMIS® physical function domain has been established. However, it has shown limited sensitivity to change and lacks utility in the clinical population (Kraft et al., 2014). The upper extremity short form has limited research to support its use independently. Moreover, focused research for its use with specific neurological conditions has not been conducted for this measure.

Therapists and other clinicians have traditionally relied on performance. Self-report measures have evolved in recent years with the Disabilities of Arm Shoulder and Hand (DASH) being the first self-report measure (Hudak et al., 1996) in hand function. Other measures such as the ABILHAND (Penta et al., 1998), the MAL (Kunkel et al., 1999), the SIS-HD (Duncan et al., 1999) and the MAM (Chen et al., 2005) were soon developed. The most recently developed self-report measures, the Neurological Quality of Life - Upper Extremity Subsection (NeuroQOL-UE) (Gershon et al., 2012) and the PROMIS® – UE (Tyser et al., 2014), were developed using the item response theory (IRT) (i.e., Rasch analysis). These two measures were developed by National Institute of Health (NIH) with its recent commitment to patient reported outcomes (Quatrano & Cruz, 2011).

All of these measures target the ICF domain of Activity and Participation. The DASH is the only measure that has attempted to address body structure and function domains by including questions related to pain, weakness, tingling, stiffness and weakness. None of the self-report measures covered other important aspects of body structure and function that address symptom burden, such as spasticity, for people with neurological conditions. Of the 7 measures, 4 were specifically developed for upper extremity function (i.e., DASH, MAM, MAL, ABILHAND) and 3 measures (i.e., SIS-HD, NeuroQOL-UE, PROMIS®-UE) include hand function within the context of a broader health assessment.

Further details of self-report measures are presented in Table 2.1 and psychometric properties in Table 2.2. The studies identified in this review that have used the self-report measures are detailed in Appendix B.

Performance-based measures. Fifteen out of the 22 hand function measures that have been researched extensively are performance-based. The details of these measures are presented in Table 2.3 and 2.4. The most commonly used performance measures were the Action Research Arm Test (ARAT) (34 studies), the Fugl-Meyer Assessment (FMA) (34 studies) and the Wolf Motor Function Test (WMFT) (11 studies). The performance-based measures are presented in the following section from older to newer measures based on the first published study found in PubMed.

Jebsen Hand Function Test (JHFT). The JHFT is also referred to as the Jebsen-Taylor Hand Function Test in literature (Beebe & Lang, 2009). The JHFT assesses broad range of unimanual hand function required for activities of daily living and is widely used in clinics across different diagnostic categories (Jebsen, Taylor, Trieschmann, Trotter, & Howard, 1969). The JHFT was reported in 4 studies out of 86 in this review. It has 7 unimanual subtests that can be completed in 15 minutes. The measure can be assembled with the detailed instructions provided in original study (Jebsen et al., 1969) however in recent years a standardized tool kit has been made available for purchase. Scoring is based on time of completion of tasks and norms have been established for dominant and non-dominant hand. The time is summed across all tasks and each task can go up to 120 seconds. The JHFT only scores time and does not include the quality of movement in its scoring criteria. The JHFT was primarily developed in orthopedic trauma where unimanual impairments were prevalent. Psychometric studies in stroke are few (Beebe & Lang, 2009; Ferreiro, dos Santos, & Conforto, 2010) and have small sample sizes. The content-related evidence for validity of scores in people with neurological conditions is lacking due to its reliance on unimanual tasks and speed with no recognition of task completion or quality of movement.

Rivermead Motor Assessment Upper Limb Section (RMA – UE). The RMA (Lincoln & Leadbitter, 1979) is a motor assessment for use in people post stroke to examine upper limb, lower limb and gross function. The RMA takes 45 minutes to administer and items are arranged in order of difficulty with 3 attempts allowed for each item. Only one study out of 86 reported using RMA in this systematic review. The RMA-UE assesses gross and fine movements and daily activity tasks using 15 items. It has unimanual and bimanual tasks. Each item is scored as a 0 if they can't perform it or a 1 if they can perform it. Even though this binary scoring system makes the scale reliable, it calls to question the utility of scores in the clinic. It is not clear how the varied range of functional levels can be represented with a 0 or 1 option for responses keeping the measure at a reasonable length of 15 items. The RMA gross motor function section can be a self-report to save therapist time (Sackley & Lincoln, 1990) however this feature has not been examined for RMA-UE. Evidence for reliability of scores for test-retest has been established for RMA-UE (r = 0.88) (Lincoln & Leadbitter, 1979) however intra-rater and inter-rater evidence for reliability of scores is lacking. Although unidimensionality (Lincoln & Leadbitter, 1979) and internal consistency (Kurtaiş et al., 2009) for the RMA-UE has been established, it lacks convergent evidence for validity of scores (Kurtaiş et al., 2009).

Motricity Index – UE Subscale (MI-UE). The MI-UE (Demeurisse, Demol, & Robaye, 1979) is a measure of movement for upper and lower extremity post stroke. This measure was reported in 3 out of 86 studies in this systematic review. This measure was developed (Demeurisse et al., 1979) for the clinic and the only equipment required is a 2.5 cm wooden cube. This was done intentionally to keep the burden low on the clinician for assembling and gathering the test items. The MI-UE has 3 items related to pinch grip using a cube for picking up and placing, elbow flexion, and shoulder abduction. No ADL items are included in this measure. Other than the advantages of ease of use with only 3 items and correlation with grip dynamometry (Collin & Wade, 1990), this measure is not comprehensive and lacks the psychometric properties for use in neurological conditions.

Action Research Arm Test (ARAT). The ARAT (Yozbatiran et al., 2008) was developed from the Carol's Upper Extremity Test (Carroll, 1965; Lyle, 1981) and is available for purchase as a standardized tool or can be constructed with the specifications provided by Yozbatiran (2008). The test forms are freely available. The ARAT is a commonly used measure and was included in 34 out of 86 studies in the systematic review.

The ARAT consists of unimanual performance tasks organized in 4 subtests of grasp, grip, pinch and gross movements. However, it has been critiqued for its lack of a theoretical model in development (Finch, Brooks, Stratford, & Mayo, 2002). This 19-item measure uses a response rating on a 4-point scale and has a standardized scoring protocol. The test takes 10 minutes to administer. All items need not be administered if the individual does not pass the first item within a subtest. This test is unique in its ability to differentiate between grasp versus grip related tasks where grip is identified as a power grip with strength. The ARAT has the most psychometric research related to stroke, with evidence for reliability of scores studied for multiple neurological diagnoses. All of the evidence for the reliability and validity of scores reported have excellent values except poor effect size in acute stroke and high floor effect in acute stroke at 14 days with 41.5 % of participants experiencing the floor effect (Lin, et al., 2009).

Fugl-Meyer Assessment – Upper Extremity Subscale (FMA-UE). The FMA-UE evaluates and measures upper extremity recovery post stroke and is the most widely used measure in stroke trials (Gladstone et al., 2002) and was reported in 34 out of 86 final studies in this systematic review. It is the only sensorimotor measure of function that includes joint pain (self-reported) measurement during passive motion, however the item has been critiqued for its representation of arthritic conditions causing joint pain versus stroke causing muscle related pain (Gladstone et al., 2002). The FMA-UE measures reflexes, movement synergies, and volitional movement. Performance is rated on a 3-point scale of none, partial and full. Items in the motor domain have been derived from Twitchell's 1951 description of the natural history of motor recovery following stroke and integrates Brunnstrom's stages of motor recovery (Gladstone et al. 2002; Poole & Whitney, 2001). Items of the FMA are intended to assess recovery within the context of the motor system. Functional tasks are not incorporated into the evaluation (Chae, Labatia, & Yang, 2003). A major criticism of the FMA is that it is a lengthy measure to administer and joint pain domain does not fit with the construct of the scale (Gladstone et al., 2002). Sometimes it takes longer than 35 minutes to complete, such as when it is administered to patients who are aphasic or severely affected (Kusoffsky, Wadell, & Nilsson, 1982; Dettmann, Linder, & Sepic, 1987). Most aspects of evidence for reliability of scores have been established for the FMA and specifically for the FMA-UE except the intra-rater evidence for reliability of scores. There have been detailed studies related to evidence for validity of scores of the measure in severity of stroke for acute, subacute, and chronic stages and is highly recommended by many of the outcome measure working groups (e.g., StrokEDGE group of American Physical Therapy Association).

Functional Test of Hemiplegic Upper Extremity (FT). The purpose of the FT is to evaluate the hemiparetic or hemiplegic arm's motor capability for function (Wilson, Baker, & Craddock, 1984). Fong et al. (2004) has adapted the FT for use in Hong Kong. The FT was reported in one out of 86 studies in this review. It consists of 17 graded activities that are arranged in 7 levels of graded degree of difficulty according to Brunnstrom's stages of stroke motor recovery. The tasks range from resisted contralateral muscle contractions to elicit associated reactions such a elbow or shoulder flexion in the impaired arm to tasks requiring a high degree of upper extremity coordination and finger dexterity of the impaired arm such as putting a light bulb into a socket held at shoulder height. The FT takes about 30 to 90 minutes to administer. The psychometric research is limited to inter-rater evidence of reliability and concurrent evidence for validity of scores only, limiting its utility for use as an outcome measure.

Box and Blocks Test (BBT). The BBT (Mathiowetz et al., 1985) is a test of unimanual gross manual dexterity with standardized equipment available for purchase. It was reported in 6 out of 86 studies in this review. It is a short test and takes about 2-5 minutes to administer. The test assesses the ability to pick up a 1-inch cube one at a time and move it across the box over a barrier within one minute. It is also a measure of endurance for repetitive tasks over time.

Multiple studies of the BBT have been conducted in neurological populations. Test-retest and inter-rater evidence for reliability of scores is excellent along with good concurrent evidence for validity of scores with the NHPT and the ARAT (Lin, Fu, et al., 2010). The BBT has also been reported as the best predictor of upper limb function at 5 weeks post-stroke (Higgins, Mayo, Desrosiers, Salbach, & Ahmed, 2005). The BBT tests only one aspect of hand function for one type of object and is best used in combination with other functional metrics for tracking hand function. Thus, it is not a comprehensive test of upper extremity function and does not include daily activity tasks.

Nine Hole Peg Test (NHPT). The NHPT (Heller et al., 1987) is test of finger dexterity administered by asking the client to take the pegs from a container one by one and place them into holes on the board as quickly as possible. Participants must then remove the pegs from the holes, one by one and replace them in the container. Scores are based on the time taken to complete the activity recorded in seconds. Alternatively the NHPT can be scored based on the number of pegs placed in 50 or 100 seconds (Sunderland, Tinson, Bradley, & Hewer, 1989). This measure was reported in 7 out of 86 studies in this review and has been validated in people with neurological conditions (Heller et al., 1987). Pinch and in-hand shifting are the only dexterity aspects assessed by this measure. Most of the psychometric studies indicate strong evidence for reliability and fair concurrent evidence for validity of scores (Erasmus et al., 2001; Sunderland et al., 1989). However, the two versions with 50 and 100 second limit have not been validated separately with 100 second being the typical method (Sunderland et al., 1989). Poor predictive evidence for validity of scores, unknown face and content-related evidence for validity of scores, ceiling effects and Minimal Clinically Important Difference (MCID) reduce the utility of this measure.

Motor Assessment Scale (MAS). The MAS (Carr et al., 1985) assesses eight areas of motor function. Three trials of each item are administered but only the best performance is recorded. Three out of the 8 areas are specifically related to the upper extremity and includes upper arm, hand movements and advanced hand activities of object manipulation. This measure was reported in 4 out of the 86 studies in this review. The MAS uses a 7-point scale of motor behavior for each area hierarchically organized from beginning to end stage of recovery. It takes about 15 minutes to administer the entire measure. The MAS targets the hand function domains of prehension, transport and non-prehensile skilled movements. Although preliminary psychometrics have been established for this measure (English, Hillier, Stiller, & Warden-Flood, 2006; Malouin, Pichard, Bonneau, Durand, & Corriveau, 1994; Tyson & DeSouza, 2004), no evidence for validity or reliability of scores exists specifically for using the upper extremity items of the MAS independently. Only one study (English et al., 2006) reported evidence for responsiveness of the arm items with 80% of subjects rating at the extremes of the scale and demonstrating poor evidence for responsiveness of the MAS - UE. Floor and ceiling effects (18% for upper arm function, 41% for hand movements and 47.5 for advanced hand activities) have been reported for the 3 upper extremity items (English et al., 2006) calling in to question the overall utility of the measure for hand function.

Sollerman Hand Function Test (SHFT). The SHFT (Sollerman & Ejeskär, 1995) is a 20-item objective measure designed to assess hand function. This measure was reported in one out of 86 studies in this review. The patient/client is given 60 seconds to complete each item and consists of a 5-point rating scale. The scale consists of items that target seven different hand grips and consists of a box with 20 items such as lock, zipper, cubes, screwdriver, etc. It targets the hand function domains of prehension and transport. This test has test-retest and inter and intrarater evidence for reliability of scores established for stroke with no evidence for validity of scores (Brogårdh & Sjölund, 2006; Weng et al., 2010). For tetraplegia and gout, there is concurrent and convergent evidence for validity of scores reported (Dalbeth et al., 2007; Sollerman & Ejeskär, 1995). There is no research regarding the measure's psychometric data for evidence of internal consistency, responsiveness, minimal clinically important difference (MCID), or floor and ceiling effects.

Arm Motor Ability Test (AMAT). The AMAT (Kopp et al., 1997) evaluates disabilities in upper extremity function in activities of daily living using a quantitative and qualitative measure. This test is unique in its breakdown of 13 ADL activities into 1 to 3 component tasks or movement segments. During the AMAT assessment, the patient/client is not aware of the analysis of each individual segment and goes through the natural flow of movement. Each task is timed and rated according to quality of movement and ability to perform each component of the task. Tasks have a 1 to 2 minute performance time limit. It is a lengthy measure to administer. It has 10-item (28 component tasks) and 9-item (without the light switch/door task) versions. It is scored on a 6-point scale from 0 (does not attempt to use affected arm) to 5 (movement appears normal). This test evaluates prehensile (grip small and large object, pinch), non-prehensile and transport functions of the hand and arm. It has adequate test retest and inter-rater evidence for reliability of scores and concurrent evidence for validity of scores (Chae, Labatia, & Yang, 2003; Kopp et al., 1997). However, only two reported studies of this measure were available and there is not sufficient study of the psychometric properties using multiple studies for this measure.

Wolf Motor Function Test (WMFT). The WMFT (Hsieh et al., 2009) was developed for the constraint induced movement therapy trials and includes a quantitative test of upper extremity motor ability through timed and functional tasks. This is a commonly used test included in 11 out of 86 studies in this systematic review. It was originally comprised of 21 items however the most commonly used version has 17 items that includes 15 function-based tasks and 2 strength based tasks. The 21-item version has 3 parts, timed tasks (tasks 1-6), functional ability (tasks 1-6, 15-21) and strength (tasks 7-14) including dynamometry for grip strength. The examiner must test the less affected extremity first and scoring is on a 6-point ordinal scale that assesses the involvement of the more affected hand from 0 (does not attempt with the involved arm) to 5 (arm does participate, movement appears to be normal). It allows for 120 seconds (2 minutes) per task. The evidence for the reliability and validity of scores of the WMFT has been studied extensively in stroke (Lang, Edwards, Birkenmeier, & Dromerick, 2008) and there is one study for TBI (Shaw et al., 2005). Overall given its strong psychometrics and its use in constraint-induced movement therapy trials, the WMFT has gained widespread use. However, since this measure was developed for research purposes, video recording of the WMFT is recommended which typically is not feasible in the clinic. Scoring accuracy is limited without video recording as a therapist might not observe all elements of the task, especially if it is done quickly or in a busy clinic.

Test Evaluant la Performance des Membres Supérieurs des Personnes Agées (TEMPA). The TEMPA (Desrosiers, Hébert, Bravo, & Dutil, 1995)was developed for the elderly population with a theoretical framework of aging hand for test development. It includes 4 unimanual and 5 bimanual tasks related to strength, coordination, dexterity and various types of prehension. The TEMPA also takes into account speed, functional rating and task analysis in scoring. Functional rating is based on a 4-point ordinal scale. Task analysis consists of range, strength, control, prehension and fine movement. Grip strength assessment using hydraulic Jamar® dynamometer accompanies the test. Normative data for young adults (20- 44 years) is included in manual

(Desrosiers, Hébert, et al., 1995). Concurrent evidence for validity of scores has been established in MS only and needs further research to obtain evidence for internal consistency, responsiveness and evidence for validity of scores in other neurological conditions (Feys, Duportail, Kos, Van Asch, & Ketelaer, 2002).

Stroke Rehab Assessment of Movement Scale (STREAM). The STREAM (Daley, Mayo, & Wood-Dauphinée, 1999) is a comprehensive measure designed for use by physical therapists for a quantitative evaluation of motor functioning after stroke. It was designed to be easy to administer in a clinical setting. It consists of upper limb (scored on a 3-point ordinal scale), lower limb (scored on a 3-point ordinal scale) and basic mobility items (scored on a 4-point ordinal scale). Extensive evidence for the reliability and validity of scores are available for this measure with overall strong psychometrics to support its use in the clinic. The upper extremity STREAM (STREAM-UE) consists of 10 items that are interspersed between the other items (Daley et al., 1999). There are no prehensile items. Movement items are assessed as the patient/client is taken in different positions from supine to sitting to standing. The lack of object handling items is a limitation of the measure.

Chedokee Arm and Hand Inventory (CAHAI). The CAHAI (Barreca, Stratford, Lambert, Masters, & Streiner, 2005) is an exclusively bimanual upper limb measure of the functional ability of the paretic arm and hand after a stroke. The CAHAI is available in 7,8, 9, or 13-item versions. The items can be assembled with instructions and training manual from the website (www.cahai.ca). The CAHAI consists of functional bimanual tasks where role of the affected hand during the functional tasks is rated on a 7-point scale of independence (similar to scoring in Functional Independence Measure). It is a newer measure with minimal testing with only 3 studies found for evidence for the reliability and validity of scores.

As occupational therapy (OT) emerged as a profession, the treatment areas were primarily work conditioning and hardening. This led to the development of hand function tools such as the Minnesota Rate of Manipulation Test (MRMT) (Drussell, 1959) and the Purdue Pegboard Test (PPT) (Rapin, Tourk, & Costa, 1966) that provided an objective assessment of hand function (e.g., turning, placing, picking tasks) needed for work environment with emphasis on time and repetition. As the role of occupational therapist included treatment of war veterans with hand injuries, strength and dexterity measures such as dynamometry (Kellor, Frost, Silberberg, Iversen, & Cummings, 1971) and functional hand measures were included. In the past 30-40 years the role of OT has expanded to include hand issues in people with disabilities related to varied orthopedic and neurological diagnoses (Radomski & Trombly, 2007). The Jebsen Hand Function Test (JHFT) is one of the oldest (Jebsen et al., 1969) measures published that has a functional upper extremity measure and continues to be the most widely used for all hand function impairments. Although bimanual measures such as the RMA (Lincoln & Leadbitter, 1979) were later developed, initially there was a heavy focus on unimanual assessment of the upper limb as evidenced by measures such as the Motricity Index – Upper Extremity Subscale (MI- UE) (Demeurisse, Demol, & Robaye, 1980), the ARAT (Lyle, 1981), the Box and Block Test (BBT) (Mathiowetz et al., 1985), and the NHPT (Mathiowetz et al., 1985). As principles of neurodevelopmental therapy were developed, measures such as the FMA – UE (1983) (Berglund & Fugl-Meyer, 1986) gained popularity with assessment of movement synergies included in the test. With increased focus on bimanual activities, measures such as the CAHAI (Barreca et al., 2005) were developed that primarily targeted bimanual function in stroke. An increasing number of measures have been developed for people with stroke, however, they have limited evidence for validity of scores across other neurological conditions.

The 15 performance measures primarily target ICF domain of activity and include everyday tasks involving upper extremities. Of these, five measures also have movement and strength measures directly addressing body function (i.e., MI-UE, FMA-UE, Functional test of hemiplegic upper extremity [FT], Motor Assessment Scale [MAS], STREAM-UE). Out of the 15 measures, there were ten measures that were specifically developed for upper extremity function (i.e., JHFT, ARAT, FT, BBT, NHPT, Sollerman Hand Function test [SHFT], Arm Motor Ability Test [AMAT], WMFT, Test d'Evaluation des Membres Supérieurs des Personnes Agées [TEMPA], CAHAI). The other five measures (i.e., FMA, STREAM, MAS, MI, RMA) evaluate hand function within the broader context of motor assessment and include lower limb and trunk related items. Only some of the upper extremity subsections like the FMA-UE were separately involved in psychometric testing, apart from their parent test.

Further details of performance-based measures are presented in Table 2.3 and psychometric properties in Table 2.4. The studies identified in this review that have reported the performance-based measures are detailed in Appendix C.

New hand function measures. This review found seven new hand function measures that have been developed in the last ten years (see Table 2.5). There were three self-report measures, the Arm Activity Measure (ArmA) (Ashford, Turner-Stokes, Siegert, & Slade, 2013), the Reaching Self-efficacy (Chen, Lewthwaite, Schweighofer, & Winstein, 2013), and the Hand Function Survey (Blennerhassett, Avery, & Carey, 2010). The ArmA (Ashford et al., 2013) was developed to separately assess the active and passive components of the affected arm. The Bilateral Arm Reaching Task (Chen et al., 2013) involves reaching self-efficacy and combines self-report and performance, wherein the participant reports confidence in reaching after performing a task. The Hand Function Survey is a self-report questionnaire that probes affected hand use in daily living tasks.

Four new performance measures were identified, the 16-Hole peg test (Hammer & Lindmark, 2010), the Extended Drawing Test (Vuillermot, Pescatore, Holper, Kiper, & Eng, 2009), the Manual Function Test (Miyamoto, Kondo, Suzukamo, Michimata, & Izumi, 2009), and the Upper Body Dressing Scale (Suzuki et al., 2008) . The 16-Hole peg test is also known as the Functional Dexterity Test and is similar to the NHPT except that the measure only requires turning over of 16 pegs using a three-jaw chuck prehension. The Extended Drawing Test, is a drawing measure. It is known as a graphonomic test that uses computerized recording of the drawings using two tools, a graphic pen with puck-like penholder and a graphic regular-grip writing pen. The Manual Function Test, developed in Japan, has 32 items and includes arm movements and manipulative activities. The Upper Body Dressing Scale for buttoned shirt is a performance assessment of donning a buttoned shirt and scores are given for each part of the task.

Among the new measures, two measures assessed bimanual function (i.e., Reaching Self-efficacy Scale and Upper Body Dressing Scale for buttoned shirt). All of the other measures focused on unimanual function and sometimes only the function of the affected hand (i.e., ArmA). The new measures of hand function attempt to reduce the time needed for assessment by developing short measures or using modern technology such as graphonomics (i.e. Extended Drawing Test). However, the measures lack assessment of technology related daily tasks performed by patients such as using a keyboard.

Discussion. There are limitations to this systematic review. The keyword search might have missed studies that did not include those keywords. Also, there might be studies not available through PubMed search engine that might have been missed in this systematic review. Further, measures developed prior to the last 10 years might still be used in clinical setting and missed in this review. The studies were research studies and may not be reflective of clinical use of these tests. The measures related to sensation were not included in this review and future studies related to hand function items in measures of sensation would likely benefit the clinician.

Numerous limitations in the existing hand function measures were identified in this review. None of the identified measures included performance and self-report in a single measure to give a comprehensive picture of patient's current hand function. There were measures like WMFT and TEMPA that combine instrumented (i.e., Jamar® Dynamometer) and performance assessment of function. There was a lack of studies to provide evidence for the reliability and validity of scores in targeted neurological conditions. The commonly used measures such as the ARAT (Yozbatiran et al., 2008), the FMA(Duncan, Propst, & Nelson, 1983), SIS-Hand (Duncan et al., 1999) and the MAL (Uswatte, Taub, Morris, Light, & Thompson, 2006) are developed and validated specific to stroke. The inability to use the measures across other diagnoses is a considerable drawback of the currently available measures. This becomes a major disadvantage in the clinic where a therapist may encounter patients/clients with multiple neurological diagnoses in a day or patients/clients who have dual diagnoses for which these measures lose evidence for validity of scores. There are measures like the DASH (Hudak et al., 1996) that have been developed specifically for people with orthopedic conditions and lack items specific to people with neurological conditions (such as spasticity) reducing its evidence for validity of scores. The currently available measures have focused exclusively on motor assessment of function. There is a lack of measures assessing sensory capacity of the hand with absence of measurement of active haptic mode of hand function. The sensory function of the hand is important for recovery and documentation of specific sensory aspects of hand function is critical. The FMA attempts to evaluate sensation however stops with assessment of passive sensory function and ignores the haptic function of hand. Furthermore, some of the measures are lengthy (e.g., FMA) or need multiple pieces of equipment set up (e.g., AMAT), making them unlikely to be utilized in a busy clinic. This review has highlighted the need to develop hand function measures with strong conceptual foundations (Baker et al., 2015), involving people with hand function deficits, and using state-of-the-art measurement techniques.

Table 2.1. Table of Evidence: Self-report Measures

Measure Author(s), year	*Construct*	*Dimensions*	*Sample item and response categories*	*Population measure developed for*	*Neurological population measure used in*	*Training and administration time*
Disabilities of Arm Shoulder and Hand (DASH) Hudak, et al., 1996 [1]	Bimanual and unimanual hand function and symptoms in daily activities, includes optional sports/performing arts or work module	ICF: Body Function, Activity and Participation Hand Function: Prehension, non-prehensile, skilled movements, transport and tactile sensing.	Rate ability to do the following activity in the past week. *"Open a tight or new jar."* Response: Difficulty rated as - 1: No, 2: Mild, 3: Moderate, 4: Severe, 5: Unable	Musculoskeletal conditions	Multiple Sclerosis, Stroke	5 to 30 minutes to administer 38 items with no training needed, manual scoring, shorter version available in 10-item format
ABILHAND Penta et al., 1998 [2]	Bimanual ability across a variety of motor impairments specific to diagnosis	ICF: Activity and Participation Hand Function: Prehension	Indicate if you can do the activity with both hands. *"Hammer a nail."* Response: N/A: Not applicable/Not an activity you attempt, 0: Not at all, 1: Only partially or with great difficulty or slowly, 2: Fully and easily	Rheumatoid arthritis	Stroke, Multiple Sclerosis, Parkinson disease	≥ 10 minutes to administer 23 items by interview with no training needed and online Rasch based automated scoring
Motor Activity Log (MAL) van der Lee et al., 1999 [3]	Bimanual and unimanual activities of daily living	ICF: Activity and Participation (Includes caregiver report) Hand Function: Prehensile, non-prehensile skilled movements, and transport	You will use two separate rating scales to describe how much and how well you use your weaker arm while you are doing specific activities like *"Open Drawer."* Response: From 0: Did not use weak arm at all, to 5: Used weak arm as good as before the stroke	Stroke, Constraint induced movement therapy trial	Stroke	20 minutes to administer 28 tasks, versions with 45, 30, 14 tasks are also available, training with reading a instruction manual, manual scoring additional time

Table 2.1. Table of Evidence: Self-report Measures (Continued)

Measure *Author(s), year*	*Construct*	*Dimensions*	*Sample item and response categories*	*Population measure was developed*	*Neurological population measure used in*	*Training and administration time*
Stroke Impact Scale - Hand Domain 3.0 (SIS-HD) Duncan et al., 1999 [4]	Bimanual and unimanual	ICF: Activity and Participation Hand Function: Prehension and non-prehensile skilled movements	In the past 2 weeks, how difficult was it to use your hand that was most affected by your stroke to... *"Turn a doorknob?"* Response: 5: Not difficult at all, 4: A little difficult, 3: Somewhat difficult, 2: Very difficult, 1: Could not do at all	Stroke	Stroke	5 minutes to administer 5 items, no training.
Manual Ability Measure-16 (MAM) Chen et al., 2005 [5]	Bimanual and unimanual	ICF: Activity and Participation Hand Function: Prehension	Please put a number to indicate how easy or how hard it is for you to do the following tasks *"Wring a towel."* Response: 4: Easy, 3: A little hard, 2: Very hard, 1: I can't do, N/A: Not Applicable	Musculo-skeletal conditions	Neurological conditions including stroke	Administration time not reported, has 16, 20, 36-item formats with manual scoring and no training needed
Neurological Quality of Life - Upper Extremity v1.0 (NeuroQOL-UE)) Carlozzi et al., 2011 [6]	Bimanual function of the upper extremities affecting quality of life	ICF: Activity and Participation Hand Function: Prehensile, non-prehensile skilled movements, and transport	Please respond to each question or statement by marking one box per row. *"Are you able to brush your teeth."* Response: 5: Without any difficulty, 4: With a little difficulty, 3: With some difficulty, 2: With much difficulty, 1: Unable to do	Traumatic Brain Injury	Diverse population, Stroke, Multiple Sclerosis, Parkinson Disease	< 2 minutes, Computer adapted testing format and paper-based versions with 8 to 10 items available
Patient Reported Outcome Measure Information System – Upper Extremity v1.2 (PROMIS-UE) Hung et al., 2011 [7]	Bimanual	ICF: Activity and Participation Hand Function: Prehensile, non-prehensile skilled movements, and transport	Please respond to each question or statement by marking one box per row. *"Are you able to reach into a high cupboard."* Response: 5: Without any difficulty, 4: With a little difficulty, 3: With some difficulty, 2: With much difficulty, 1: Unable to do	Diverse populations	Neurologic populations	Administration time not reported, 16 items with online scoring available

Note: ICF: International Classification of Functioning, Disability and Health; [1] Hudak, Amadio, & Bombardier, 1996; [2] Penta et al., 1998; [3] van der Lee et al., 1999 ; [4] Duncan et al., 1999 ; [6] Carlozzi, Tulsky, & Kisala, 2011 ; [7] Hung, Clegg, Greene, & Saltzman, 2011

Table 2.2. Table of Evidence: Self-report Measures Psychometric Properties

Measure Author(s), year	*Evidence for reliability of scores*	*Evidence for validity of scores*	*Evidence for responsiveness*
Disabilities of Arm Shoulder and Hand (DASH) Hudak et al., 1996 [1]	Test-retest: In Athletes, ICC = 0.536 [2]; Proximal Humerus Fractures, ICC 2,1 = 0.928 (0.860 to 0.963) [3]; Total elbow arthroplasty, ICC = 0.96 [4].	Criterion-related evidence for validity of scores: Neck pain - concurrent evidence for validity of scores with VAS = 0.55, NDI = 0.82-0.84 [5] Construct-related evidence for validity of scores: Neck pain - convergent validity of scores - SF-12 physical component = 0.62 [6], Carpometacarpal arthroplasty – discriminant evidence for validity of scores –SF 36 physical component = -0.49 [7] Content and Face validity evidence: N/E for neurological conditions. Poor for stroke [8]. Established for musculoskeletal conditions [1].	Shoulder, arm and hand = 1.92 [6], Neck pain = 1.38 [6], Total elbow arthroplasty – Effect size = 0.56, Sensitivity = 0.59, Specificity = 0.71 [4]. Japanese: Wrist disorders – Effect size = 1.20, Standard response mean =1.30 [9] MDC: Athletes – 10 [10], Proximal humeral fractures – 16.1 [3] MCID: Athletes – 10 [10], Post-op surgery for UE – 19 (15 to 23) for patients reporting "much better" [11] Ceiling effects: Elbow disorders 1-6% [4], Proximal humeral fractures [3] - 7% ceiling. Floor effect was found to 0% in elbow disorders [4]
ABILHAND Penta et al., 1998 [12]	Stroke: Separation Rasch evidence for reliability of scores = 0.90 [13]. Stroke: DIF analysis – uniformity established [13]. Stroke, MS, SCI, PD, cerebellar ataxia, healthy – Cronbach's alpha = 0.99 [14]	Stroke - All items fit Rasch model, 23 items defined manual ability, item distribution = 1.72 to -2.18 logits [13]. Content-related evidence for validity of scores established [13]. Factor analysis – score explained 84% of variance [14] Criterion-related evidence for validity of scores: Stroke- Concurrent evidence for validity of scores with Jamar (r=0.481), BBT (r=0.481) [14] Construct-related evidence for validity of scores: Stroke: Known group – Unaffected limb with Jamar (r = 0.242), BBT (r = 0.248) [13]. Affected limb with Jamar r = 0.562, BBT r = 0.598 [13]. Differences between tetraparesis, hemiparesis, other (MS, PD, cerebellar ataxia) & controls [14]	Average ability of healthy controls = 89 (SE=8) vs patients with stroke, MS, PD, Cerebellar ataxia, SCI, PD group = 63 (SE=17) [14], Stroke: Sensitivity = 92% and Specificity = 80% [14] MDC: 0.13 logits [15] Stroke - Measurement error in center of scale = 0.36 logits [15]
Motor Activity Log (MAL) van der Lee et al., 1999 [16]	Test-Retest evidence for reliability of scores: Stroke: MAL – Amount of use (AOU) r = 0.70 to 0.85, MAL – Quality of movement (QOM) r = 0.61 to 0.71 [17]. Cronbach's alpha MAL AOU = 0.88, QOM = 0.91 [17].	Stroke: Concurrent evidence for validity of scores MAL 28 QOM and SIS-HD r = 0.72 [18], MAL 28 QOM and accelerometry r = 0.52 [17], MAL 28 and ARAT r = 0.63 [18]	Responsiveness ratio AOU = 1.9, QOM = 2.0 with Spearman rho = 0.16 to 0.22. MAL 14 QOM participant scale responsiveness ratio > 3 [17]

Table 2.2. Table of Evidence: Self-report Measures Psychometric Properties (Continued)

Measure *Author(s), year*	*Evidence for reliability of scores*	*Evidence for validity of scores*	*Evidence for responsiveness*
Stroke Impact Scale - Hand Domain 3.0 (SIS-HD) Duncan et al., 1999 [19]	Test-Retest evidence for reliability of scores: SIS Acute stroke - ICC 0.70 to 0.92 [20] Cronbach's alpha = 0.83 to 0.90 [20], SIS HD ICC = 0.95 [21] for SIS in Acute stroke	Concurrent evidence for validity of scores of SIS-HD and Fugl-Meyer Assessment-Upper extremity, r = 0.81 [19], Criterion-related evidence for validity of scores SIS- HD rho = 0.51 to 0.68 [20] Construct-related evidence for validity of scores: Convergent evidence for validity of scores - SIS and Burden of Stroke Scale r = -0.83 [22]. Discriminant evidence for validity of scores for various degrees of stroke severity [19]	SIS-HD medium evidence for responsiveness SRM = 0.52 [23]. SIS- HD SEM = 9.4 [23] MDC: SIS- HD = 25.9 [24] MCID: SIS-HD = 17.8 [24] Floor effect 2% in minor stroke, 40.2 % in moderate stroke, ceiling effect 14.6% in minor stroke, 4.9 % in moderate stroke [19]
Manual Ability Measure-16 (MAM) Chen et al., 2005 [25]	Test-Retest, Inter-rater, Intra-rater evidence for reliability of scores: N/A Person separation 0.93, Item separation 0.99. Variance explained by measure was 78.4% [25] in Rheumatoid Arthritis, Osteoarthritis, Carpal tunnel syndrome, tenosynovitis, hand injuries	Criterion-related evidence for validity of scores: Moderate correlation with LIFEware physical function r = 0.51, SF-12 moderate correlation r = 0.44 [26] Construct-related evidence for validity of scores: Item discrimination greater than 0.4 [25]. MAM 16: Known group: Rasch person evidence for reliability of scores = 0.83. Fit statistics not greater than 1.4. Item evidence for reliability of scores = 0.98 Content-related evidence for validity of scores: DIF 14 items - Single construct, no misfitting items [25]	Not Available
Neurological Quality of Life – Upper Extremity Section v1.0 (Fine Motor, ADL) (NeuroQOL-UE) Carlozzi et al., 2011 [27]	Test-Retest evidence for reliability of scores: ICC 0.73 to 0.94 [28, 29] Cronbach's alpha 0.85 to 0.97 [28] for NeuroQOL	Criterion-related evidence for validity of scores: Predictive: Addressed for entire set of items [29] Construct-related evidence for validity of scores: Short forms correlated to full item bank (r = 0.82 to 0.96), Short forms able to differentiate people with 0, 1-2, 3 reported diagnoses. Face validity evidence: Patient telephone interviews, focus groups [27].	Not Available

Table 2.2. Table of Evidence: Self-report Measures Psychometric Properties (Continued)

Measure *Author(s), year*	*Evidence for reliability of scores*	*Evidence for validity of scores*	*Evidence for responsiveness*
Patient Reported Outcome Measure Information System – Upper Extremity Section v1.2 (PROMIS-UE) Hung et al., 2011 [30]	Test-Retest evidence for reliability of scores: ≥ 0.80 [32] Unidimensionality to physical function [31,30] in Diverse population	Criterion-related evidence for validity of scores: Global physical health scale, r=0.52 [31]. Correlation with DASH (r = 0.726) [33] Construct-related evidence for validity of scores: PROMIS® -UE and mobility able to explain 35% of variance in overall PROMIS® scores [31]. PROMIS® -UE correlated with overall scores (r = 0.64). DIF established for patients with upper extremity and mobility issues [31].	Diverse population: Addressed, more responsive than mobility items [31]. Floor and ceiling effects: < 1%. Diverse population: Skewness and kurtosis for the estimated upper extremity scores in the sample were -1.70 (-0.88) and 2.19 (-0.07). SEM: 3.3 to 5.9

Note: ADL: Activities of Daily Living; BBT: Box and Block Test; DIF: Differential Item Functioning; ICC: Intraclass Correlation Coefficient; MCID: Minimally Clinically Important Difference; MDC: Minimal Detectable Change; MID: Minimally important difference; MS: Multiple Sclerosis; N/A: Not Applicable; NDI: Neck Disability Index; OA: Osteoarthritis; PD: Parkinson disease; RA: Rheumatoid arthritis; SCI: Spinal Cord Injury; SEM: Standard Error of Measurement; SF: Short Form Health Survey; UE: Upper Extremity; VAS: Visual Analog Scale; [1] Hudak et al., 1996 ; [2] Alberta et al., 2010 ; [3] Slobogean, Noonan, Famuyide, & O'Brien, 2011 ; [4] Angst et al., 2012 [5] Mehta et al., 2010 ;[6] Huisstede, Feleus, Bierma-Zeinstra, Verhaar, & Koes, 2009 ; [7] MacDermid, Wessel, Humphrey, Ross, & Roth, 2007 ;[8] Lannin et al., 2010 ;[9] Imaeda et al., 2010 ; [10] Hsu, Nacke, Park, Sennett, & Huffman, 2010 ; [11] Gummesson, Atroshi, & Ekdahl, 2003 ; [12] Penta et al., 1998 ; [13] Penta et al., 2001 ;[14] Simone et al., 2011 ; [15] Penta, 2015 ;[16] van der Lee et al., 1999 ;[17] van der Lee et al., 2004 ; [18] Uswatte et al., 2005 ;[19] Duncan et al., 1999 ; [20] Duncan et al., 2002 ;[21] Carod-Artal et al., 2009 ; [22] Doyle et al., 2007 ;[23] Keh-Chung Lin et al., 2010 ; [24] K Lin, Fu, et al., 2010 ; [25] Chen et al., 2005 ; [26] Chen & Bode, 2010 ; [27] Carlozzi et al., 2011 ; [28] Cella et al., 2012 ; [29] Cella, 2015 ; [30] Hung, Clegg, Greene, & Saltzman, 2011 ; [31] Hays et al., 2013 ; [32] Broderick, Schneider, Junghaenel, Schwartz, & Stone, 2013 ; [33] Tyser et al., 2014

Table 2.3. Table of Evidence: Performance-based Measures

Measure Author(s), year	*Construct*	*Dimensions*	*Sample item and response categories*	*Population measure developed for*	*Neurological population measure used in*	*Training and administration time*
Jebsen Hand Function Test (JHFT) Jebsen et al., 1969 [1]	Unimanual activities of daily living	ICF: Activity Hand Function: Prehension, transport	Checkers: *Place your left hand on the table please. When I say 'Go', use your left hand to stack these checkers on the board in front of you as fast as you can like this one on top of the other (demonstrate). You may begin with any checker. Do you understand? Ready? Go.* Response: Timed	Spina Bifida	Stroke, Brain injury, Spinal Cord Injury	15 minutes for 7 items with assembling of materials needed and training needed but not specified in the manual
Rivermead Movement Assessment – Upper Extremity (RMA-UE) Lincoln et al., 1979 [2]	Bimanual and unimanual movements and activities of daily living	ICF: Activity Hand Function: Prehension, non-prehensile skilled movements, transport	Lying, protract shoulder girdle with arm in elevation. Arm may be supported. Response: 1 if patient can perform activity, 0 if he cannot	Stroke	Stroke	45 minutes for total test and less for 15 arm items with assembly of materials needed
Motricity Index – Upper Extremity (MI-UE) Sunderland et al., 1989 [3]	Unimanual movements	ICF: Body structure and function Hand Function: Prehension, transport	Elbow Flexion to 90 degrees so that the arm touches the shoulder. Response: 14 points are given if movement is seen with the elbow out and the arm horizontal	Stroke	Stroke	Less than 5 minutes for 3 items, only one cube needed, no training
Action Research Arm Test (ARAT) Lyle, 1981 [4]	Unimanual movements	ICF: Activity Hand Function: Prehension, non-prehensile skilled movements, transport	Pick up 10 cm block. Response: 0 – No movement, 3 – Normal movement	Stroke	Stroke, Multiple Sclerosis, Traumatic Brain Injury	5 to 20 minutes for 19 movements in 4 subscales, assembly of materials needed, training needed

Table 2.3. Table of Evidence: Performance-based Measures (Continued)

Measure *Author(s), year*	*Construct*	*Dimensions*	*Sample item and response categories*	*Population measure developed for*	*Neurological population measure used in*	*Training and administration time*
Fugl-Meyer Assessment – Upper Extremity (FMA-UE) Duncan et al., 1983 [5]	Unimanual active and passive movements, reflexes, speed, sensation and joint pain	ICF: Body structure and function Hand Function: Prehension, non-prehensile, skilled movements, tactile sensing	Pronation-supination, elbow at 90 degrees and shoulder at 0 degrees Response: 0=cannot perform, 1=performs partially and 2=performs fully Response options are tailored to the item	Stroke	Stroke	20 minutes for 7 sections on upper extremity.
Functional Test for the Hemiplegic Upper Extremity (FT) Wilson et al., 1984 [6]	Bimanual and unimanual activities of daily living, associated reactions and movements	ICF: Body structure and function, activity Hand Function: Prehension, non-prehensile skilled movements, transport	Level 2B. Hand into lap. Response: Pass or fail based on ability to complete task within 3 minutes except Level 1 and 2A.	Stroke	Stroke	30 minutes for 17 graded activities
Box and Blocks Test Mathiowetz et al., 1985 [7]	Unimanual pinch	ICF: Activity Hand Function: Prehension	Number of blocks moved from one partition to the other in 1 minute	Musculo-skeletal conditions	Multiple Sclerosis, Stroke, TBI, Neuromuscular diseases, Spinal Cord Injury	1 minute, one task
Motor Assessment Scale – Upper extremity (MAS-UE) Carr et al., 1985 [8]	Unimanual activities of daily living and movements	ICF: Body structure and function, activity Hand Function: Prehension, non-prehensile skilled movements, transport	Upper arm, hand and advanced hand activities with graded levels from 0 to 6 with different tasks within each grade level. Advanced hand activities grade 6: Hold a comb and comb hair at back of head (shoulder externally rotated, abducted at 90 and head erect).	Stroke	Stroke	15 minutes for 8 activities of entire measure, time needed for materials and scoring

Table 2.3. Table of Evidence: Performance-based Measures (Continued)

Measure Author(s), year	*Construct*	*Dimensions*	*Sample item and response categories*	*Population measure developed for*	*Neurological population measure used in*	*Training and administration time*
Nine Hole Peg Test (NHPT) Heller et al., 1987 [9]	Unimanual pinch	ICF: Activity Hand Function: Prehension	Seconds to completion of pegs placed on board	Musculoskeletal conditions	Brain Injury, Stroke, Parkinson disease	Varies with patient's ability, approximately 10 minutes
Sollerman Hand Function Test (SHFT) Sollerman et al., 1995 [10]	Bimanual and unimanual 8 grip related tasks	ICF: Activity Hand Function: Prehension, transport	Fold paper, put into envelope. Response: 5-point rating scale from 0 – Task cannot be performed at all, to 4 – Task completed without any difficulty within 20 seconds and with the prescribed hand-grip of normal quality. Subjective estimate of hand function on a 10 cm line from no hand function to full hand function	Musculoskeletal conditions, tetraplegia	Tetraplegia, Stroke	20 minutes for 20 subtests
Test Evaluant la Performance des Membres supérieurs des Personnes âgées (TEMPA) Desrosiers et al, 1995 [11]	Bimanual and unimanual activities of daily living	ICF: Activity Hand Function: Prehension, transport	Open a jar and take a spoonful of coffee. Grip strength with dynamometer. Response: Speed, task analysis and function rating. Task analysis – range, strength, control, prehension, fine movement. Function rating on a 4-point ordinal scale	Geriatrics (60 years and over)	Multiple Sclerosis, Stroke	9 tasks, demands on administration and scoring unknown.
Arm Motor Ability Test (AMAT) Kopp et al., 1997 [12]	Bimanual and unimanual activities of daily living	ICF: Activity Hand Function: Prehension, non-prehensile skilled movements, transport	Cut meat Response: Functional rating and quality of movement use a 6-point ordinal scale from 0 to 5. Performance emphasized not speed. Combined function and quality rating into one scale is also available.	Stroke, Constraint Induced Movement Therapy Trials	Stroke	30 to 40 minutes, materials to assemble and lengthy to complete the 28 tasks with 1 to 2 minutes limit per task. 9 or 10 item version may be available.

Table 2.3. Table of Evidence: Performance-based Measures (Continued)

Measure Author(s), year	*Construct*	*Dimensions*	*Sample item and response categories*	*Population measure developed for*	*Neurological population measure used in*	*Training and administration time*
Stroke Rehabilitation Assessment of Movement Scale (STREAM) Daley et al., 1999 [13]	Unimanual movements	ICF: Body structure and function Hand Function: Non-prehensile skilled movements, transport	*Shrug your shoulders as high as you can.* Response: 4 point ordinal scale, 0 – unable to perform the test activity through any appreciable range, 3 – able to complete the activity independently with a greatly normal movement pattern without an aid	Stroke	Stroke	20 minutes entire test, 10 upper limb items with no prehensile items
Wolf Motor Function Test (WMFT) Wolf et al., 2001 [14]	Bimanual and unimanual activities of daily living	ICF: Activity Hand Function: Prehension, non-prehensile skilled movements, transport	*Place your forearm on the table as quickly as you can. Do it just like this (examiner demonstrates). At the end of the movement, your forearm and hand should be touching the surface of the table. Do this as quickly as you can. (repeat instructions) Do you have any questions? Ready, set, go.* Response: Time score for speed and Functional ability scale is a 6-point ordinal scale where 0-does not attempt with the involved arm and 5 – arm does participate/movement appears to be normal	Stroke, Constraint Induced Movement Therapy Trials	Stroke, Traumatic Brain Injury	60 minutes for each upper extremity tested on 15 items. Materials, training with a DVD and scoring time needed.
Chedokee Arm and Hand Inventory (CAHAI) Barreca et al., 2004 [15]	Bimanual activities of daily living	ICF: Activity Hand Function: Prehension, non-prehensile skilled movements, transport	*Draw a straight line the length of the ruler using both of your hands* Response: 7 point activity scale from 1 – Total assistance to 7 – Complete independence	Stroke	Stroke, Brain injury	30 minutes Time for materials, training and scoring needed.

Note: ICF: International Classification of Functioning, Disability and Health; [1] Jebsen et al., 1969 ; [2] Lincoln & Leadbitter, 1979 ; [3] Sunderland et al., 1989 ; [4] Lyle, 1981 ; [5] Duncan et al., 1983 ; [6] Wilson et al., 1984 ; [7] Mathiowetz et al., 1985 ; [8] Carr et al., 1985 ; [9] Heller et al., 1987 ; [10] Sollerman & Ejeskär, 1995 ; [11] Desrosiers, Hébert, et al., 1995 ; [12] Kopp et al., 1997 ; [13] Daley et al., 1999 ; [14] Wolf et al., 2001 ; [15] Barreca et al., 2004

Table 2.4. Table of Evidence: Performance-based Measures Psychometric Properties

Measure *Author(s), year*	*Evidence for reliability of scores*	*Evidence for validity of scores*	*Evidence for responsiveness*
Jebsen Hand Function Test (JHFT) Jebsen et al., 1969 [1]	Test-retest evidence for reliability of scores: Stroke: r = 0.92 [2], Tetraplegia: Dominant hand r = 0.91 - 0.99 [1] dominant hand 'writing' r = 0.67, non-dominant hand r = 0.78 – 0.92, non-dominant hand 'feeding' r = 0.60 and non-dominant hand 'large objects' r = 0.67. No practice effect was observed [1]. Inter-rater evidence for reliability of scores: Stroke: Portuguese version - ICC = 1.0 [3] Healthy ICC=0.82 -1.00 [4] Intra-rater evidence for reliability of scores: Stroke: Portuguese version - ICC = 0.997 [3] Cronbach's alpha = 0.924. Stroke: Portuguese version of JHFT 'writing' alpha = 0.884, Pearson's r (item to total score) for 'small objects' r = 0.657 [3]	Criterion-related evidence for validity of scores: Concurrent evidence for validity of scores: Stroke: Correlation with grip strength r = 0.79 – 0.81, pinch strength r = 0.60 = 0.79, ARAT r = 0.87 – 0.95, NHPT r 0.84 – 0.97, SIS-HD r 0.61 – 0.83 [2] Correlation with MHQ subtests for RA, OA, Fracture r = 0.10 – 0.41 [5]. CTS correlation for ADL r = 0.68, function r = 0.69 [5]. Construct-related evidence for validity of scores: Discriminant evidence for validity of scores Positive = change in MHQ scores > 20, negative = change in MHQ score $\leq$ 20. ROC area under curve for RA = 0.52, OA = 0.58, CTS = 0.66, Fracture = 0.59. [5]	Stroke: 1-3 months post stroke, effect size = 0.69 and between 1-6 months post stroke = 0.73 [2] Effect size, CTS = 0.05, Fracture = 0.35, RA = 0.47, OA = 0.67 [5].
Rivermead Movement Assessment – Upper Extremity (RMA-UE) Lincoln et al., 1979 [6]	Stroke: Test-Retest Evidence for reliability of scores: RMA-UE r = 0.88 [6]. Inter-rater: Significant difference in scoring across rater as compared to variability between patients. Revised scoring produced for RMA-UE – testing of revised version not done [6]. Stroke: RMA-UE Cronbach's alpha = 0.95, ICC = 0.93 [7]	Content-related evidence for validity of scores: Unidimensionality with Guttman scale [6]. Criterion-related evidence for validity of scores: RMA concurrent with Barthel Index across each assessment period initial (r=0.84), 1 month (r=0.78), and one year (r=0.63) [8]. Predictive evidence for validity of scores – low RMA scores at 6 weeks post stroke predicted with poor prognosis to ambulate [9]. Construct-related evidence for validity of scores: Convergent evidence for validity of scores RMA-UE with MI r = 0.74 [9], RMA with FIM cognition r = 0.52 [10], RMA-UE with FIM r = 0.386 – 0.483 [7]. Alternative forms convergent evidence for validity of scores with verbal and performance methods of completion r = 0.98 [11], Known group validity of scores with RMA score and infarct size at initial r = -0.52, 1 month r = -0.47, and 1 year r = -0.53 [8].	All 3 sections are sensitive to change, arm section least responsive when compared to FIM. RMA ES = 0.38, SRM= 0.60 [7] MCID: $\pm$ 3 points likely to represent a clinically relevant change in functional level [12]. Large ceiling on RMA gross motor subscale when compared to HIMAT [13]. Floor and ceiling not established for RMA-UE.

Table 2.4. Table of Evidence: Performance-based Measures Psychometric Properties (Continued)

Measure *Author(s), year*	*Evidence for reliability of scores*	*Evidence for validity of scores*	*Evidence for responsiveness*
Motricity Index – Upper Extremity (MI-UE) Sunderland et al., 1989 [14]	Stroke: Inter-rater evidence for reliability of scores: MI-UE = 0.88 (Spearman's rho) [15]. Stroke: Cronbach's alpha = 0.968 [15].	Stroke: Criterion-related evidence for validity of scores: Concurrent with dynamometry of UE r = 0.89 [16]. RMA-UE with MI Spearman's rho 6 weeks = 0.76, 12 weeks = 0.73, 18 weeks = 0.74. Predictive for acute stroke at 1, 3, 6 months post stroke [15]. Predictive MI combined with trunk control test at 6 weeks predicted failure to walk by 18 weeks [15].	Response mean = 0.96, Effect size = 0.27 [17]. Stroke: Floor/Ceiling effects: 57% had had measureable pinch grip within first 3 weeks of a stroke, 2% had normal pinch grip [14]. Sensitivity: Early change with acute stroke can be detected [14].
Action Research Arm Test (ARAT) Lyle, 1981 [18]	Test-retest evidence for reliability of scores: ICC = 0.894 to 0.976, rho = 0.897 to 0.976 [19], r = 0.98 [18] Inter-rater evidence for reliability of scores: Acute stroke: ICC = 0.92 [20], Chronic stroke: ICC = 0.995 [21]. Chronic and acute stroke, MS, TBI: ICC = 0.998 [22], Chronic stroke: Inter-rater: ICC = 0.9986, rho = 0.96 [23] Intra-rater evidence for reliability of scores: Chronic stroke: ICC = 0.989 [21]	Criterion-related evidence for validity of scores: Chronic stroke: Concurrent with FMA-UE [21] r = 0.94 [23] and MAS [21]. Construct-related evidence for validity of scores: Convergent with FMA motor domain (rho = 0.925), MI (rho = 0.811) and BBT (rho = 0.951) [19]	Chronic Stroke: Responsiveness ratio = 2.03, 1.7 points with summing items [21]. Acute stroke: Effect size between 14 and 180 days = 0.79 [24] Effect size at 1 and 3 months post stroke = 0.55, 1 and 6 months post stroke = 0.63 [25], 0 and 90 days post stroke 1.390 [26]. MID, MDC: N/E Chronic stroke: MCID = 10% of total (5.7 points), Acute stroke: MCID Dominant side = 12 points (ES = 0.78), Non-dominant side = 17 points (ES = 1.10) [21]. Normative Data: Established [21,25] Acute stroke: Floor effects for raw scores < 3 and ceiling for > 54 [20], Floor effects 14 days post stroke = 41.5 % [24], ceiling effects 180 days post stroke = 22.6 % [24]. Acute stroke: Cronbach's alpha 0.985 [20].

Table 2.4. Table of Evidence: Performance-based Measures Psychometric Properties (Continued)

Measure Author(s), year	*Evidence for reliability of scores*	*Evidence for validity of scores*	*Evidence for responsiveness*
Fugl-Meyer Assessment – Upper Extremity (FMA-UE) Duncan et al., 1983 [27]	Stroke, MS, TBI: Test-Retest Evidence for reliability of scores: FMA Total ICC = 0.97 [28], Chronic stroke: Inter-rater evidence for reliability of scores: FMA total r = 0.98 to 0.99, FMA-UE r = 0.98 to 0.995 [27]. Stroke: Cronbach's alpha = 0.94 to 0.98 across 4 administrations [29]. FMA-UE to FMA r = 0.97 [30]	Criterion-related evidence for validity of scores: Stroke: Concurrent evidence for validity of scores: FMA and MAS r = 0.96 [31], Poor FMA and MAS sitting balance r = - 0.10 [31], Predictive evidence for validity of scores: Motor and sensory FMA scores 5 days post stroke predictive of motor recovery 6 months post-stroke [32]. Construct-related evidence for validity of scores: Acute stroke: Convergent with FMA and Barthel Index (r = 0.86 – 0.89) [33] and FIM (r = 0.63) [34]. Known group evidence for validity of scores: FMA distinguished between dependent, partly dependent and independent levels of self-care [35]. Chronic Stroke: Convergent with FMA-UE and Barthel Index (r = 0.75) [36], FMA and ARAT (rho = 0.73) [37], WMFT time (rho = 0.76) [37], FIM-motor (rho = 0.49) [37]. Content-related evidence for validity of scores: Acute stroke: FMA-UE modified with 3 reflex items removed, adequate fit using Rasch Analysis except for hook grasp item at admission and 6 months post stroke [38]. Chronic stroke: Unidimensional items within the FMA-UE [39]. Face validity evidence: Established and heavily weighted for UE [40].	Chronic Stroke: Responsiveness ratio = 0.41 [41]. FMA – Sensation subscale SRM at 14 - 180 days = 0.67 [29]. Shortened FMA ES = 0.53, FMA ES = 0.045 [42] FMA-UE: MDC = 5.2 points [43], MCID = 10 points = 1.5 change in discharge FIM [44]. Floor/Ceiling effects Stroke: Possible ceiling effects with hand [39] Stroke: Cronbach's alpha = 0.94 to 0.98 across 4 administrations [29]. FMA-UE to FMA r = 0.97 [30]
Functional Test for the Hemiplegic Upper Extremity (FT) Wilson et al., 1984 [45]	Stroke: Inter-rater evidence for reliability of scores, rho = 0.976 [45]. Hong Kong version, rho = 0.93 [46].	Criterion-related evidence for validity of scores: Concurrent with FMA r=0.88 and FMA-UE subscores r=0.88. Correlation with FIM self care, r=0.46 [45] Construct-related evidence for validity of scores: FT scores accounted for 87% of variance in six other test scores [45]	Not available

Table 2.4. Table of Evidence: Performance-based Measures Psychometric Properties (Continued)

Measure Author(s), year	*Evidence for reliability of scores*	*Evidence for validity of scores*	*Evidence for responsiveness*
Box and Blocks Test (BBT) Mathiowetz et al., 1985 [47]	Acute and Chronic Stroke: Test-retest reliability: More affected hand r = 0.98, Less affected hand r = 0.93 [48]. Upper limb impairment: Right hand: ICC = 0.90, Left hand: ICC = 0.89 [49]. Stroke, MS, TBI: ICC = 0.96, Spastic Hemiplegia: ICC = 0.95 [50] Inter-rater reliability: Healthy: Left hand r = 0.99, Right hand = 1.00 [47], Healthy: ICC = 0.80 [51], Stroke, MS, TBI: ICC 0.99 [52], Spastic hemiplegia: r = 0.95 [50], Fibromyalgia ICC = 0.85 [51].	Stroke: Criterion-related evidence for validity of scores: Concurrent post treatment with NHPT (r = - 0.71), ARAT (r = 0.64), FMA (r = 0.35), SIS (r = 0.52), MAL-AOU (r = 0.49), MAL-QOM (r = 0.52). Adequate with FMA, MAL and SIS [53]. Predictive evidence for validity of scores: At 1 and 5 weeks post-stroke, BBT was best predictor of upper limb function [54]. Construct-related evidence for validity of scores: UE Paresis: Convergent with ARAT (r = 0.95), FMA (r = 0.92), Hemispheric stroke scale (r = -0.67), passive joint motion/joint pain subscale (r = 0.43), Modified Barthel Index (r = 0.04) [52]. Brain injury associated paresis (stroke, anoxic brain damage, TBI, SCI): Convergent with resistance to passive movement (r = - 0.680) [52].	Acute Stroke: 1 and 4 weeks post-stroke SRM: 0.8 [54], 0 and 3 weeks SRM = 0.74 [53]. MDC: Acute and Chronic stroke: 5.5 blocks per minute (18% percent change) [48]; Spastic Hemiplegia: 4 blocks per minute for 2 week training [50].
Motor Assessment Scale (MAS) Carr et al., 1985 [55]	Test-Retest Evidence for reliability of scores: Acute and chronic stroke: r = 0.87 to 1.00 [55] Inter-rater evidence for reliability of scores: Acute stroke 87% agreement [55]	Criterion-related evidence for validity of scores: Acute stroke: Concurrent with modified FMA (rho = 0.96) [56], similar MAS to FMA items rho = 0.65 to 0.93 except FMA balance items (rho = 0.12) [56]. Construct-related evidence for validity of scores: Acute stroke: Convergent: MAS and functional balance test, sitting arm raise (r = 0.33) and sitting forward reach (r = 0.54) [57], Mobility Scale for Acute Stroke mobility items convergent with MAS r = 0.89 [58].	Stroke: ES: 0.36 to 0.5 [59], Subjects with no change on the measure 44.4 to 63.9% [59], MAS arm items – 80% of subjects rated at the extremes of the scales [59]. Significant differences in mean scores 71 days after rehabilitation for stroke [60]. Floor/Ceiling effects: Stroke: Floor effects on admission for upper arm function (18%), hand movements (41%) and advanced hand activities (47.5%). Floor effects on discharge for hand movements (27.9%) and advanced hand activities (36.1%). Stroke: Ceiling effects on admission for upper arm function (19.7%), hand movements (18.0%). Ceiling effects on discharge for upper arm function (41%), hand movements (36.1%), advanced hand activities (16.4%) [59].

Table 2.4. Table of Evidence: Performance-based Measures Psychometric Properties (Continued)

Measure Author(s), year	*Evidence for reliability of scores*	*Evidence for validity of scores*	*Evidence for responsiveness*
Nine Hole Peg Test (NHPT) Heller et al., 1987 [61]	Test-retest evidence for reliability of scores: Acute and chronic stroke: ICC = 0.85 [62], without spasticity ICC = 0.86 [62], with hand spasticity ICC = 0.64 [62], PD: Dominant hand ICC = 0.88 [63], Non-dominant hand ICC = 0.91 [63], Healthy: Right hand ICC = 0.95 [64], Left hand ICC = 0.92 [64], MS: Better hand rank correlation coefficient = 0.923, Worse hand = 0.862 [65]. Inter-rater evidence for reliability of scores: Healthy: Right hand r = 0.984 [66], Left hand r = 0.993 [66], MS: MS composite with NHPT part of it ICC = 0.95 [67], Stroke: r=0.75 to 0.99 [61]. Intra-rater evidence for reliability of scores: MS: MS composite with NHPT part of it, ICC = 0.97 [67], Stroke: r = 0.68 to 0.99 [61]	Criterion-related evidence for validity of scores: Concurrent evidence for validity of scores: Stroke: Poor concurrent with Frenchay Arm test – 27% cases improperly grouped [68]. Stroke: Concurrent with SIS-HD pre-treatment (rho = 0.58), post-treatment (rho = -0.66) [69], BBT and ARAT pre-treatment (rho = -0.55 to -0.80) and post-treatment (rho = -0.57 to -0.71) [68]. FMA and MAL pre-treatment (rho = -0.16 to -0.27), post-treatment (rho = -0.18 to -0.33) [68]. Healthy: Concurrent with Purdue Pegboard Test (rho = -0.74 to -0.75) [64]. Predictive evidence for validity of scores: Poor at 1 month [68]. Construct-related evidence for validity of scores: Stroke: Convergent with Motricity Index r = 0.82.	Stroke: ES at 1 and 3 months post stroke = 0.52, 1 and 6 months post stroke = 0.66 [70], SRM = 0.64 to 0.79 [69] MDC: Stroke = 32.8 seconds (Percent change = 54%) [62], PD: MDC Dominant hand = 2.6 seconds, Non-dominant hand =1.3 seconds [63]. Floor effect: Stroke: 100 second cut-off score - 20% with minimum score at discharge [71], 50 second cut-off score – poor floor effects [68]
Sollerman Hand Function Test (SHFT) Sollerman et al., 1995 [72]	Test-Retest Evidence for reliability of scores: Chronic Stroke ICC = 0.96 - 0.98 [73] Chronic Burns: Inter-rater evidence for reliability of scores ICC = 0.98, Intra-rater evidence for reliability of scores ICC = 0.98 [74]. Chronic Stroke: Inter-rater ICC = 0.96 - 0.98, Intra-rater evidence for reliability of scores ICC = 0.96 - 0.99 [73]	Criterion-related evidence for validity of scores: Tetraplegia: Concurrent with subjective hand function on visual analog scale (r = 0.68), Disability Rating Scale for insurance purposes in Sweden for hand dysfunction (r = 0.88) [72]. Gout: Concurrent with hand tophus joint (r = 0.59) [75] Construct-related evidence for validity of scores: Tetraplegia: Convergent with Motor Capacity Scale (r = 0.959), Gout: Convergent with Grip strength (r = 0.731) [75], DASH (r = 0.887) [75].	MDC: Chronic Burns = 6.7 – 6.9 [74] Chronic Burns: SEM = 2.6 [74]

Table 2.4. Table of Evidence: Performance-based Measures Psychometric Properties (Continued)

Measure Author(s), year	*Evidence for reliability of scores*	*Evidence for validity of scores*	*Evidence for responsiveness*
Test Evaluant la Performance des Membres supérieurs des Personnes âgées (TEMPA) Desrosiers et al, 1995 [76]	Test-Retest Evidence for reliability of scores: Older adults: ICC 0.70 to 1.0 [76] Inter-rater evidence for reliability of scores: Older adults: Non-simultaneous inter-raters ICC = 0.68 to 1.0, simultaneous inter-raters = 0.96 to 1.0.	Criterion-related evidence for validity of scores: MS: Concurrent with JHFT r = 0.56 to 0.87 [77], NHPT r = 0.79 to 0.9, FIM r = 0.44 and ADL self-questionnaire r = 0.61 [77]. Construct-related evidence for validity of scores: Convergent with functional independence in basic personal care (rho = 0.74)	Not reported
Arm Motor Ability Test (AMAT) Kopp et al., 1997 [78]	Test-retest evidence for reliability of scores: Stroke: ICC 0.93 - 0.99 [78] Inter-rater evidence for reliability of scores: Stroke: ICC=0.95-0.99 [78]	Criterion-related evidence for validity of scores: Stroke: Concurrent with FMA = 0.92-0.94 [79] Construct-related evidence for validity of scores: Stroke: Convergent with MI-UE r = 0.45 - 0.61 [78]	Stroke: 1 versus 2-week change detected in subacute stroke for mild to moderate movement deficits [78].
Stroke Rehabilitation Assessment of Movement Scale (STREAM) Daley et al., 1999 [80]	Test-Retest Evidence for reliability of scores: Chronic Stroke: Mobility subscale ICC = 0.96 (0.93 to 0.98) [81] Chronic Stroke: Inter-rater evidence for reliability of scores UE = 0.994, Total = 0.995. Intra-rater UE = 0.963, Total = 0.999 [80]. Chronic Stroke: Cronbach's alpha overall score = 0.984, limb subscales = 0.979 [80]	Criterion-related evidence for validity of scores: Acute stroke: Concurrent with STREAM - UE with Box and Block affected UE initial = 0.78 and 3 months = 0.76. Unaffected UE initial = 0.31 and 3 months = 0.31 [82]. Construct-related evidence for validity of scores: Acute stroke: Convergent with Barthel Index 14 days = 0.80 at 14 days and 180 days = 0.76. Predictive with Barthel Index 0.54 at 14 days and 0.81 at 90 days. Face validity evidence: 20 physical therapists reviewed and refined the measure [80]	Acute stroke: 5 weeks post-stroke STREAM SRM = 0.98, STREAM-UE= 0.75 [83]. Admission and 48 hours of discharge STREAM- UE ES = 0.38, SRM = 0.78 (Change score 3.3) [84]. 14 and 180 days post-stroke STREAM SRM = 1.65 [85] Chronic Stroke: Mobility subscale of STREAM: SEM = 1.5 points [81]. MDC: Chronic Stroke: SRM = 4.2 points [80]. Acute stroke UE STREAM = 2.8 points (14.0%) [84]. SRM = 4.2 for mobility subscale of STREAM [81] MCID: Stroke: STREAM UE = 2.2 points [84]. Floor effects: Acute stroke: STREAM-UE at admission = 26%, discharge = 8% [84]. Ceiling effects: Acute Stroke: STREAM-UE at admission = 20%, discharge = 40% [84]. After 180 days ceiling effect for 16.3% [85]

Table 2.4. Table of Evidence: Performance-based Measures Psychometric Properties (Continued)

Measure Author(s), year	*Evidence for reliability of scores*	*Evidence for validity of scores*	*Evidence for responsiveness*
Wolf Motor Function Test (WMFT) Wolf et al., 2001 [86]	Test-Retest evidence for reliability of scores: Chronic TBI: ICC 0.97 (0.89 to 0.97) [87], Chronic stroke r = 0.95, ICC = 0.97 [87]. Inter-rater evidence for reliability of scores: Chronic stroke: ICC 0.93 to 0.99 [88]	Criterion-related evidence for validity of scores: Stroke: Concurrent with FMA-UE r = -0.57 to -0.88 [86] Construct-related evidence for validity of scores: Stroke: Known group evidence for validity of scores of scores: WMFT score for individuals without impairment were significantly higher than clients with stroke for both dominant and non-dominant hand [86]	SRM: Acute stroke: Time = 0.38, functional ability scale (FAS) = 1.30 [89] SEM: 0.2 seconds [90], Evidence for MDC: Chronic stroke: Time: 0.7 seconds, FAS: 0.1 points [90]. MCID: Acute stroke: FAS dominant side affected =1.0 (17% change), time dominant side affected = -19.0 seconds (16% change) [91] Evidence for Ceiling effects: Chronic stroke: 5 to 17% of patients [92]
Chedokee Arm and Hand Inventory (CAHAI) Barreca et al., 2004 [93]	Test-retest evidence for reliability of scores: UE paralysis: ICC=0.96 [94] Inter-rater evidence for reliability of scores: UE paralysis: ICC=0.98 [94] Cronbach's alpha = 0.98 [93]	Criterion-related evidence for validity of scores: Concurrent evidence with ARAT, CMSA, ROC curves CAHAI - 7= 0.95, CMSA = 0.76, ARAT = 0.88 [95] Construct-related evidence for validity of scores: Known groups evidence: ROC curves for CAHAI – 13 = 0.86, CAHAI – 9 = 0.82 greater than ARAT = 0.72 [95]. Content/Face validity evidence: Reported for stroke [94]	Stroke: ROC curves for CAHAI – 13 = 0.86, CAHAI – 9 = 0.82 were greater than ARAT = 0.72, demonstrating sensitivity of CAHAI to change [94]. UE paralysis MDC (90) = 6.3 points [94]

Note: ADL: Activities of Daily Living; CMSA: Chedokee McMaster Stroke Assessment; CTS: Carpal Tunnel Syndrome; ES: Effect Size; FIM: Functional Independence Measure; ICC: Intraclass Correlation Coefficient; MCID: Minimal Clinically Important Difference; MDC: Minimal Detectable Change; MHQ: Michigan Hand Questionnnaire; OA: Osteoarthritis; RA: Rheumatoid Arthritis; ROC: Receiver Operating Curve; SCI: Spinal Cord Injury; SEM: Standard Error of Measurement; SRM: Standard Response Mean; TBI: Traumatic Brain Injury; UE: Upper Extremity; [1] Jebsen et al., 1969 ; [2] Beebe & Lang, 2009 ; [3] Ferreiro et al., 2010 ; [4] Hackel, Wolfe, Bang, & Canfield, 1992 ; [5] Sears & Chung, 2010 ; [6] Lincoln & Leadbitter, 1979 ; [7] Kurtaiş et al., 2009 ; [8] Endres, Nyary, Banhidi, & Deak, 1990 ; [9] Collin & Wade, 1990 ; [10] Soyuer & Soyuer, 2005 ; [11] Sackley & Lincoln, 1990 ;[12] Collen, Wade, & Bradshaw, 1990 ; [13] Williams, Robertson, Greenwood, Goldie, & Morris, 2006 ;[14] Sunderland et al., 1989 ; [15] Collin & Wade, 1990 ;[16] Bohannon, 1999 ;[17] Vos-Vromans, de Bie, Erdmann, & van Meeteren, 2005 ; [18] Lyle, 1981 ; [19] Platz et al., 2005 ; [20] Nijland et al., 2010 ; [21] van der Lee, Beckerman, Lankhorst, & Bouter, 2001 ; [22] Platz et al., 2005 ;[23] Yozbatiran et al., 2008 ; [24] Lin, Wu, Liu, Chen, & Hsu, 2009 ;[25] Beebe & Lang, 2009 ; [26] Lang, Wagner, Dromerick, & Edwards, 2006 ; [27] Duncan et al., 1983 ; [28] Platz et al., 2005 ; [29] Lin, Hsueh, Sheu, & Hsieh, 2004 ;[30] Wood-Dauphinee, Williams, & Shapiro, 1990 ;[31] Malouin et al., 1994 ; [32] Duncan, Goldstein, Matchar, Divine, & Feussner, 1992 ; [33] Mao, Hsueh, Tang, Sheu, & Hsieh, 2002 ; [34] Shelton, Volpe, & Reding, 2001 ;[35] Bernspång, Asplund, Eriksson, & Fugl-Meyer, 1987 ;[36] Dettmann, Linder, & Sepic, 1987 ; [37] Hsieh et al., 2009 ; [38] Woodbury et al., 2008 ; [39] Crow & Harmeling-van der Wel, 2008 ; [40] Gladstone et al., 2002 ; [41] van der Lee et al., 2001 ;[42] Hsueh et al., 2008 ;[43] Wagner, Rhodes, & Patten, 2008 ; [44] Shelton, Volpe, & Reding, 2001 ; [45] Wilson et al., 1984 ;[46] Fong et al., 2004 ; [47] Mathiowetz et al., 1985 ;[48] Chen, Chen, Hsueh, Huang, & Hsieh, 2009 ; [49] Desrosiers, Bravo, Hébert, & Dutil, 1995 ; [50] Siebers, Oberg, & Skargren, 2010 ;[51] Canny, Thompson, & Wheeler, 2009 ; [52] Platz et al., 2005 ; [53] Lin, Chuang, Wu, Hsieh, & Chang, 2010 ; [54] Higgins et al., 2005 ; [55] Carr et al., 1985 ;[56] Malouin et al., 1994 ; [57] Tyson & DeSouza, 2004 ;[58] Simondson, Goldie, & Greenwood, 2003 ; [59] English et al., 2006 ; [60] Dean & Mackey, 1992 ; [61] Heller et al., 1987 ; [62] Chen et al., 2009 ; [63] Earhart et al., 2011 ; [64] Wang et al., 2011 ;[65] Erasmus et al., 2001 ; [66] Oxford Grice et al., 2003 ; [67] Cohen et al., 2000 ; [68] Sunderland et al., 1989 ; [69] Lin, Chuang, et al., 2010 ;[70] Beebe & Lang, 2009 ; [71] Jacob-Lloyd & Dunn, 2005 ; [72] Sollerman & Ejeskär, 1995 ; [73] Brogårdh & Sjölund, 2006 ;[74] Weng et al., 2010 ; [75] Dalbeth et al., 2007 ; [76] Desrosiers, Hébert, et al., 1995 ; [77] Feys et al., 2002 ; [78] Kopp et al., 1997a ; [79] Chae et al., 2003 ; [80] Daley et al., 1999 ; [81] Chen et al., 2007 ; [82] Ahmed et al., 2003 ;[83] Higgins et al., 2005 ; [84] Hsueh et al., 2008 ; [85] Hsueh, Wang, Sheu, & Hsieh, 2003 ; [86] Wolf et al., 2001 ; [87] Shaw et al., 2005 ; [88] Morris et al., 2001 ; [89] Hsieh et al., 2009 ;[90] Fritz, Blanton, Uswatte, Taub, & Wolf, 2009 ; [91] Lang et al., 2008 ; [92] Nijland et al., 2010 ; [93] Barreca et al., 2004 ; [94] Barreca et al., 2005 ; [95] Barreca et al., 2006

Table 2.5. New Measures of Hand and Arm Function Identified in the Review

Measure	*International Classification of Functioning, Disability and Health category*	*Laterality*
Self-report		
Arm Activity Measure (ArmA)	Activity and Participation	Unimanual
Hand Function Survey	Activity	Unimanual
Reaching self efficacy (RSE), camera recording during testing	Body structure and function	Bimanual
Performance		
16-Hole Peg Test	Activity	Unimanual
Extended Drawing Test (computerized)	Activity	Unimanual
Manual Function test	Activity	Unimanual
Upper Body Dressing Scale for buttoned shirt	Activity	Bimanual

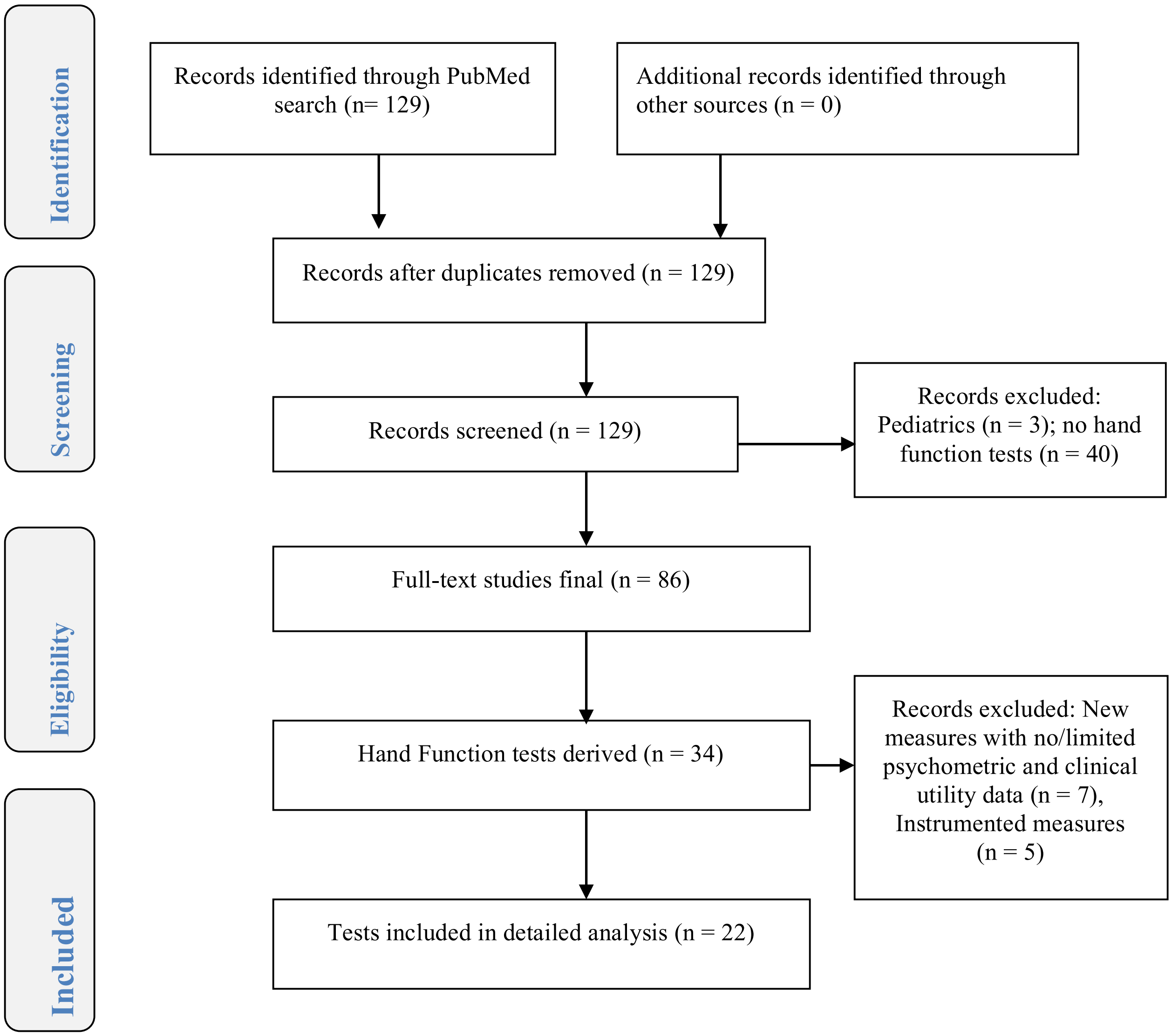


Figure 2.1. PRISMA Flow Diagram

Chapter 3: Construct Definition and Expert Panel Study

Abstract

Background and Purpose: A new measure – Hand and Arm Function Measure (HAFM) - was developed for assessing upper extremity function in daily activities in people with neurological conditions. This clinical measure has a combination of self-reported and performance-based items. The conceptual rehabilitation framework of International Classification of Functioning, Disability and Health (ICF) and the measurement framework of evidence-centered design (ECD) shaped the development of this measure. An expert panel study was undertaken to shed light on the definition of the construct, item development, and item writing.

Method**:** The construct of 'hand and arm function in daily activities' is defined for the purpose of this measure using a construct map and the relationships to other variables are described using the nomological network. Experts were involved in two stages: Construct and item bank development and item writing and modification to establish face and content-related evidence for validity of scores. The four experts included people in clinical practice of physical and occupational therapy in neurological conditions, teaching, research, and measurement. The construct and item bank development stage involved a qualitative descriptive study with the focus group of experts. The item writing and modification was conducted after patient/client focus groups and was in the form of an open-ended questionnaire. The identified items were examined, suggestions for item addition or deletion were made by the experts, and their feedback was incorporated.

Results and conclusion: The construct of 'hand and arm function in daily activities' was defined, item bank development started and item modifications were considered. The construct

map, nomological net and expert panel study informed the development of the first iteration of the self-reported and performance-based sections of the Hand and Arm Function Measure.

Construct Definition

Introduction. Hand and arm function has varied perspectives for developing an operational definition. The perspectives differ mainly in the extent of proximal function involved in the conceptualization. The definition could include reflexes as defined in Fugl-Meyer Assessment Scale – Upper Extremity Section (FMA-UE) (Duncan et al., 1983). It could include more proximal body functions like scapular function in Stroke Rehabilitation Assessment of Movement – Upper Extremity Section (STREAM-UE) (Daley et al., 1999). It can take into account shoulder function as defined in the measure DASH (Hudak et al., 1996). It could take into account grip strength as defined in Jamar® dynamometer (Mathiowetz et al., 1985). It could take into account only wrist and hand function as defined in the measure Nine Hole Peg Test (NHPT) (Oxford Grice et al., 2003). There is no published operational definition of the construct for most of the hand and arm function measures in literature, except for Motor Activity Log (MAL) that claims to measure 'real world hand use' (Uswatte, et al., 2005).

Defining the construct operationally provides a framework for analysis of the concept to be measured (Ketterlin-Geller, Yovanoff, Jung, Liu, & Geller, 2013). Some authors doubt the utility of creating a general definition of the construct (Nesselroade, Gerstorf, Hardy, & Ram, 2007). The concern relates to the individual differences among people that cannot be measured with a generic construct. For example, low level of the trait of hand and arm function could mean poor ability to do tasks needing strength for one individual or high pain in the wrist for another individual. This concern cannot be alleviated easily, however defining all the possible combinations leading to a low level and high level of trait can clarify the understanding and use of the scores. Even though, hand and arm function appears to be a simple construct, it has many aspects that can be measured (e.g., grip strength, sensation, dexterity, etc.) making it a complex construct to operationally define. Thus, the construct definition needs to take into consideration the wide range of dimensions to be measured in developing an operational definition (Sadler, 2003).

Objective. The objective of this study was to define and describe the construct of 'hand and arm function in daily activities' using a construct map and the relationships to other variables using a nomological network.

Definition. Good or high trait of hand and arm function was defined as the ability of a person to participate in a wide range of daily activity tasks with the use of their hands and arms with optimal speed and quality, without experiencing interference by body structure and function problems when the environmental conditions remain constant. The guiding principles for development of the measure include: a) combination of self-reported and performance-based items, b) feasibility to administer in a clinical setting, c) evidence for reliability of scores for measuring high and low trait of hand function in neurological conditions, and d) sensitive to change.

This construct has two dimensions, gross and fine movements in daily activities based on current literature (Lehman et al., 2011). 'Gross movements in daily activities' dimension is defined as proximal arm movements such as transport, reaching, carrying, and pushing. Transport and reach is typically performed in combination with prehension. Pushing and carrying may not accompany prehension. 'Fine movements in daily activities' dimension was defined as prehensile movements like grip, and non-prehensile skilled movements like turning and pointing. Fine movements in daily activities for the purpose of this measure includes sensory function that is essential for hand function. The sensory task that is used in the performance-based measure involves fine movements of small objects in a sock without visual input and thus falls within the fine movements dimension. The hand functions were defined with Jones and Lederman's classification system (Jones & Lederman, 2006) described previously in Chapter 2.

Although, the hypothesized dimensions in the measure are gross and fine movements, it is possible that the dimensions of body structure and function, and activity and participation based on International Classification of Functioning, Disability and Health (ICF) (WHO, 2001) may emerge predominantly. ICF framework was the theoretical framework selected for development of this measure. The acknowledgement of this possibility is important during factor analysis studies in the later stages of development of the measure when the dimensionality will be explored statistically. As the measure is based on ICF and the rating of symptoms within the self-report section may not fit neatly into gross and fine movement categories, further investigation of the domains may be needed.

ICF framework. The World Health Organization's International Classification of Functioning, Disability and Health (ICF) (WHO, 2001) provides an common taxonomy of health-related domains. The ICF core set for hand conditions was used as a broad framework for the self-reported measure (Coenan, 2015). An European Delphi Study has taken this core set and identified the existing impairment and activity measures related to hand and arm function that will help assess the aspects in the core set (van de Ven-Stevens et al., 2016). This approach however sets up a clinician for using multiple measures to assess aspects of the hand and arm core set as opposed to the current measure that attempts to unify the measurement process to one single measure. The core set provides lists of categories that are relevant for hand function and disability to support the application of ICF in practice. It takes into account all the possible areas within body structure, body function, activity and participation that are influenced by hand function problems. Although, hand function is affected by environmental and personal contexts as outlined in the core set, the Hand and Arm Function Measure assumes these factors are constant when measuring body function, activity and participation. Table 3.1 provides the construct schema in detail with the representation of ICF categories, hand function domains and corresponding items in the Hand and Arm Function Measure.

Evidence Centered Design (ECD) Framework. The ECD framework is a measurement framework that enables systematic study of the development of the new measure (Mislevy et al., 2003). The categories in ECD framework are discussed previously in Chapter 1. Note that these categories can overlap and multiple iterations of the processes can exist. The layers of ECD and their application for the Hand and Arm Function Measure are summarized in Table 3.2 and subsequently described.

Domain analysis. The content, concepts and representations of hand and arm function are studied in domain analysis and have undergone revisions over the course of this study. For example, it was expected for the name of the measure to reflect the purpose. The original name for the Hand and Arm Function Measure was 'fine motor function measure.' As the work to accumulate evidence for the background related to the various components of hand function developed, the domain of fine motor no longer adequately reflected the purpose. Fine motor control (e.g., dexterity related movements such as picking up coins) was also analyzed theoretically to have minimal value in the absence of gross motor (e.g., arm, forearm and hand combined movements such as reaching for a glass on a shelf) information to have meaning for the interpretation of scores. Moreover, the term 'motor' underrepresented the varied domains underlying movements such as sensation, control, and coordination. The revised domain thus emerged to be one of 'real world hand use' and since it is not feasible to 'observe' real world performance, this was revised to 'hand use.' The need to change this further to reflect 'arm and hand use' was considered to accurately reflect what aspects of hand function the measure supported. Expert and participant (individuals with neurological conditions) input further revised this to fit the ICF model and the name of the construct was changed to its current form of 'hand and arm function in daily activities.' The evidence for reliability of scores for quantifying hand use was a concern expressed by expert participants. The ambiguity of the term 'use' with the implication of 'use for what purpose' was expressed by participants with neurological conditions. Moreover, the term 'upper extremity' although better fits the name of the measure for the therapists, it is not the terminology that the patients/clients can identify with. Thus, this patient/client centered measure used easily understood terms of 'hand and arm' to covey the gross and fine movement aspects to the patients/clients.

A nomological network or net is a schematic description of the factors that affect the construct 'hand and arm function in daily activities' that explains the relationships of the aspects within and outside of the construct. It has been used in measurement for many decades (Cronbach & Meehl, 1955). The nomological network for HAFM is shown in Figure 3.1 and is organized with the ICF framework as the overarching framework with the domains under each category shown with their relationship with hand and arm function. The solid lined text boxes in Figure 3.1, indicate the continuum from the health conditions that this construct addresses that affect the person's body function resulting in hand and arm function problems that manifest in the activity and participation domains tested in this measure (Lemmens et al., 2012). The dashed lined boxes indicate the other body structure and function, environment, and personal factors that can affect the construct. Other health conditions affecting the construct and changes in the demands of the task are other considerations not shown in the figure that can affect the factors in this nomological network. Although evidence such as that provided by Lemmens et al., (2012) to support each relationship in this network is desirable, deeper research is lacking in this field.

The construct map is shown in Figure 3.2. It explains the organization of the construct to reflect high to low function and how that will correspond to the change in scores and level of difficulty on the measure. High scores on the measure indicate better function or high level of the trait of hand and arm function. Low scores on the measure indicate poor hand and arm function.

Domain modeling. The Toulmin diagrams for assessment arguments (Mislevy et al., 2003) and 'design patterns' (Mislevy, 2011) are the ways in which domain modeling was represented. Table 3.2 gives the design pattern attributes with the corresponding terminology in Messick's Assesment Argument and Toulmin Assessment Argument with the application to the HAFM. The Toulmin Assessment Arguments can be demonstrated with a diagram for Hand and Arm Function Measure as shown in Figure 3.3. The warrant is the evidence for the claim made by the measure. Identifying these warrants and claims helps clarify the intentions of the HAFM.

The construct definition has undergone changes over the course of this study. These changes are summarized in Table 3.3. The changes not only impact the design of the assessment but they also impact the studies designed for testing evidence for validity of scores of the measure. The major changes have been with the change in focus from purely fine motor to fine and gross movements. The target population was also clearly defined and its effect on the design of the evidence for validity of scores was included to reduce construct under-representation (e.g., by including wider population) and construct irrelevant variance (e.g., by measuring aspects not related to the construct).

Conceptual Assessment Framework. This process organizes the purpose, practical or financial limits in order to create the framework of the test. This framework is organized into three perspectives (Messick, 1994), student (patient/client), task and evidence.

a) Student (patient/client) model. In the student model, the claims of patient/client's knowledge and skills that can be made by the interpretation of test scores are addressed. For the Hand and Arm Function Measure, we have evidence from prior literature that disorders can cause impairments in the sub-domains of strength, sensation, mobility, stability, pain and coordination which can cause subsequent difficulty in function such as in reach, grasp, release, in-hand manipulation, unilateral and bilateral control (Radomski & Trombly, 2007). These patient/client-related variables can vary and contribute to the information we can get about the patient/client from the scores.

b) Task model. The tasks that should elicit the behaviors are addressed in this model. The environment that influences the participant is studied as it affects the performance on the measure. The evidence behind using object manipulation tasks for the Hand and Arm Function Measure falls into this category (Hsieh et al., 2009; Jebsen, Taylor, Trieschmann, Trotter, & Howard, 1969). The concept of introducing gradation of tasks in the Hand and Arm Function Measure based on the function and control sub-domains is an element of the task model. A change in domain modeling to include reach, grasp, release, in-hand manipulation, unilateral and bilateral control also impacted the task model which now has to consider a more broader collection of items to address these sub-domains.

c) Evidence model. An evidence model connects the student (patient/client) and task models. The behaviors that should reveal the construct are translated into scores. This has two components, evaluation and measurement. The evaluation component addresses the scoring rules, for example, manual versus computer based. Also, decisions such as scoring rubrics are made within the evaluation component. For the Hand and Arm Function Measure, the observed tasks have a scoring form that relates to completion of task levels and compensations used during task, this rubric can help generate scores. For the self-report items, the current version has a paper and pencil single version format with response choices related to interference of hand and arm function issues or symptoms with daily tasks. The measurement component includes decisions about using statistical models such as classical test theory models or the more recent item response theory (IRT) models. The selection of classical test theory model for Hand and Arm Function Measure was made considering the moderate sample size estimates for the neurological population for the pilot study. The scores in the two sections of the hand and arm function measures would be combined. Cluster analysis would be desired to identify normal, mild, moderate and severe impairment categories. According to the Code of Fair Testing Practices in Education (Joint Committee on Testing Practices, 2004) informing test takers about results is considered a standard in education. However, clinics are high stakes testing situations where client's adverse responses or indifference to information about the results can be difficult to handle. Many clinicians choose not to share the results with clients unless specifically requested by clients. This practice needs to change in order to be more transparent with our clients and share the results in a meaningful way. Thus, this new measure incorporates this belief and ways to inform the clients about the results.

Assessment implementation. This involves the actual constructing of the test and its operational elements. For the Hand and Arm Function Measure, the self-report items were put together in a format with ease of reading, reading level and clarity of instructions under consideration. The writing of items is another process that sought expert input. The observation task rules were generated however the test materials are still under development in this phase. Moreover, as domain analysis and modeling become more formed and clear, the actual test construction undergoes changes at every stage. For example, the response choices for the self-report items in the prior version were ‘none of the time,’ ‘some of the time,’ ‘most of the time,’ and ‘all the time.’ However, after putting the items together in the form of a questionnaire, it was noted that some people never engaged in those tasks that were asked in the questionnaire. The responses were then revised to be able to mark ‘N/A’ to capture this information that could have potentially introduced irrelevant information in the test. Also, the response choices were changed to ‘not at all,’ ‘a little,’ and ‘quite a lot.’ The total scores were reflected to accommodate N/A and missed responses. To develop a good quality measure, there is a need to be constantly in the process of reducing construct-irrelevant information and under-representation that refines the construct definition and reiterates the process multiple times.

Assessment delivery. This is the process where the patient/client engages in the test. The four-process delivery architecture, discussed earlier in this paper, guides this process. Testing the Hand and Arm Function Measure with patients/clients would include the process of activity selection, presentation, collecting data, and accumulating evidence. The purpose of the test is always revisited in the process of making these decisions. The test results can be claimed to be useful within the studied diagnostic groups alone. Functional groups of mild, moderate and severe impairments based on scores may emerge that may be useful to distinguish one condition from another based on the spread of scores. If any claims will be made based on this information, the construct definition would then need to be reworked to address those uses.

Expert Panel Study

Introduction. In order to develop a new measure of hand and arm function in neurological conditions, the evidence-centered design framework was adopted (Mislevy et al., 2003). According to the framework, the development of an item-bank is an important first stage along with the previously discussed construct definition. Experts in the field are the stakeholders who will use the measure and based on their experience have opinion about the development and use of a hand and arm function measure. Involving experts early on provides analysis of the construct, its operational definition and the development of an item bank.

Purpose. A qualitative descriptive study was undertaken to gain expert input on the definition of the construct of hand and arm function in daily activities, item development, and item writing. This was part of the first phase of development of the measure and consisted of four stages (refer to Figure 3.4 for schematic representation of the stages). Experts were involved in two stages: Stage I: Construct and item bank development and Stage III: item writing and modification to establish face and content-related evidence for validity of scores. Stage II and IV involved patient/client participants in focus groups and cognitive interviews respectively, and will be discussed in Chapter 4.

Method. A qualitative descriptive study research methodology using face-to-face semi-structured format in a focus group was used to obtain in-depth data from experts in the field of occupational therapy, physical therapy and measurement. The pragmatic interview methodology was selected over surveys in order to help the participants understand the construct to be studied; to bring out hidden beliefs about assessment of arm and hand function, identify professional descriptors currently used in therapy and measurement; and to gather ideas about potential items through group validation. Human subjects approval was obtained from University of Washington

Human Subject Division. A semi-structured interview guide (see Table 3.4) was developed based on literature regarding hand and arm function, the need to define the construct, and for item development. After informed consent, the focus group discussion was conducted at the University of Washington with four expert participants and two researchers (NG[1] and DK[2]). The interviews were audio-recorded with the informed consent of the participants and transcribed verbatim. A detailed account of the interview was maintained by the researcher.

Analysis. Data immersion, coding, category creation and thematic analysis were used to find patterns of meaning across data (Thomas, 2006; Schmidt, 2004). The researchers used an inductive approach to derive themes by interpreting the raw data. Coding categories and subcategories were allocated by two researchers independently. The researcher's comments were coded separate from the expert comments to prevent bias during coding. The researcher's notes after the session were not coded. The group discussion and item modification feedback was pooled for analysis. The data management software Dedoose version 6.2 was used for managing transcripts, codes and themes.

Results. The study recruited experts from the field of occupational therapy, physical therapy and measurement. To ensure that views from teaching and research in therapy were also represented, the participants were purposefully sampled. The sample size was limited by budget and participants were invited via a written and verbal invitation. Four expert participants were recruited with their demographic information provided in Table 3.5. There were two occupational therapists (OT), one physical therapist (PT) and one measurement expert. Of the two OTs, one had mainly clinical experience with adult neurological population, the other OT had a combination of teaching and clinical experience with children. The PT had clinical,

[1] NG: Namrata Grampurohit
[2] DK: Deborah Kartin

research and teaching background with Parkinson disease population as the main area of interest with an added focus on fine motor function research. All three therapists have had experience with assessment and treatment of people with stroke. The measurement specialist had scale development, teaching and research background with research experience in rehabilitation. All expert participants approached agreed to participate in the study. All expert participants expressed interest in the measurement of arm and hand function and actively participated in the focus group. The audio-recordings were transcribed verbatim. There was overall good agreement (89.35%) and the items of disagreement (10.65%) were resolved by discussion to reduce researcher bias during thematic development phase. The disagreement was calculated by identifying excerpts marked and discussed by the researchers. Both researchers agreed on the final coding system and accepted it as being representative of the data. Seven key themes emerged and will be discussed in detail in the order in which they appear in the semi-structured interview guide.

Hand function versus hand use. The expert participants tried to differentiate between the two terms 'use' versus 'function.' They were able to identify hand function as related to 'properties' of the hand and use as specific skills that include function and bilateral use. However, they were not able to clearly point out a difference in their usage in the clinic. Moreover, they viewed both the terms under the activity and participation domain of the ICF (International Classification of Functioning, Disability and Health) with synonymous usage in clinic.

> *I pretty much use them synonymously, use and function. Just because I think I look at the framework that the ICF provides and I just put function or classify function in their activities domain. (E4)*

In view of the suggestion to ICF related terminology and unclear distinction between use and function, the name of the construct was changed from 'hand use' to 'hand function.'

Simulated versus real life tasks. Although the expert participants agreed that there is a difference between simulated versus real life tasks for assessment, the pros and cons of using both type of tasks were discussed. There was a desire to observe real life tasks because they are automatic and functional.

> *I think that it's always nice to see them do real life things because it's more automatic. So, that's kind of how I see that as more functional and the simulated is where they may not have ever done that before or it's not as automatic to them. (E1).*

This theme validated the need of the self-report measure and it was evident that the focus would need to be real tasks that are common in daily activities. It was also discussed that compensation plays a role in real life tasks to get the tasks done efficiently and cannot be ignored or eliminated from the measure. However, experts in research emphasized the utility of simulated measures.

> *I agree completely that doing the real thing is the real thing and cannot be substituted. But, I also think that sometimes a simulated test in a research environment does give you a lot of information that may be missed otherwise with a functional test. Like for example, if you were to look at force profiles with say an instrument that has force sensors on it and it simulates say jar opening. Especially, in the population that I am familiar with which is early Parkinson's disease. In their function, people may not show as many deficits especially if they are on medications. But with an instrumented device, you may see deficits in force profile that may not be evident on clinical examination. (E4)*

Thus, the experts felt the need for sensitive measurement in the case of mild impairments, which validated the need for the performance section with simulated items that might be able to capture mild forms of impairment due to measurement of speed. However, since this measure is not designed for pure research but mainly for use in clinic, the instrumented measurement was not deemed necessary.

Compensations and adaptations. The experts agreed that people compensate in various ways such as stabilizing their postural muscles to bring about arm movements. Thus, the rating scale for the performance-based assessment should incorporate a category that includes compensations. Since adaptive devices and splints can be additionally applied to the upper extremity, the patient/client information section at the beginning of the self-report and performance-based measure would allow patients or therapist to enter this information. However, it was noted that some compensations can cause damage in the long run and it is important to identify them in assessment.

> *Or you feel that in the long term it may cause some other deficit like shoulder pain or something like that which you need to prevent at the right time. (E1)*

The performance-based section of the new measure needs has to have provision for therapists to note such abnormal movements within the structure of the measure. Since these aberrations can be different for each individual and different for the various diagnoses, the rating scale cannot be specified to score lower for these movements.

There was a concern voiced about patient's perception of the performance-based assessment.

> *I think one of the items that missed in a lot of test is that we don't get an indication of how easy or difficult it is for the patient. (E4)*

This concern indicated a need for patient-report in the performance-based assessment. It could not be incorporated for each item on the measure, as it would add to the length of administration of the measure. However, it does bring out the need for patient's perspective on the items administered in the test. The way this concept was incorporated in the test is having the patient/client choose one item from the measure that provided them just right challenge and they felt was important to them to track their progress in therapy.

Conditions. The experts noted that the hand function deficits do differ with the diagnosis and some conditions like Parkinson disease could have specific symptoms like tremors and medication related effects that should be considered if the test is to be valid for use in people with Parkinson disease. There was a discussion about external cues needed for people with PD during intervention, the issues related to progression of the disease and presence of ataxic arm movements. The more frequency of excerpts was found for the three PD symptoms of tremors, slow movements and rigidity and these were added as symptoms to the self-report measure. The issue of good function on medication (also known as 'on – phase') and poor function off medication (also known as 'off – phase') is specific to PD.

So external cues might be specific to Parkinson's, even on and off medications, it's a huge difference. It is necessary to know how long it has been since their last dose of medication. (E4)

The details about medications were included in the patient/client information section of both self-report and performance-based sections of the measure.

For stroke, experts had comments about learned non-use, hand preference, tone, using the strong hand to help the weak hand, and problems with releasing objects.

And also for stroke you might be changing your hand preference. (E1)

The change in hand function was captured in patient/client information section with question about current and prior hand preference. Tone was added in the symptom list as spasticity and rigidity symptoms. For multiple sclerosis, the experts agreed that it presented with a varied range of symptoms that are mostly bilateral. For TBI, the experts mentioned that there were more cognitive deficits, bilateral involvement and may not have the presentation of hemiplegia. The idea of providing more cues during testing was discussed in relation to TBI and it was suggested that although cues would be needed for some patients they do serve as a form of assistance and should be treated accordingly in the measure.

> *Breaking it down into smaller steps. But again that problem with simulated versus real world comes in right, if you always have the test incorporate a cue like put the pen on a shelf, they may not have a shelf in their life, they may have to just lift something up. They may not have an external cue all the time. (E3)*

People with neurological conditions have other associated orthopedic issues.

> *So, my mother had rheumatoid arthritis and I think about that. (E2)*

It seemed important to have people list all their diagnoses in the patient/client information section at the beginning of the test for the therapist to have information about other diagnoses that could affect their hand function.

Items. Fifty-six items were suggested by experts from experience with patients with neurological conditions. These items were added to the item bank to be combined with items suggested by patient/client participants in focus groups for a total set of 144 items in the item bank. A subset of the items in the item bank were to be administered to patient/client participants for cognitive interviews and a method for item extraction was needed. For the purpose of item development, a strategy for triangulation of results from expert participants and patient/client participants was developed. The frequency of occurrence of each item among patients and experts was combined with the frequency of co-occurrence of each item for all groups. The items were then ranked by their total occurrence score and top one-third of the items that had total occurrence score of 6 or higher were selected for further development. Items were later added based on theoretical constructs to be addressed.

> *Paper and pencil, puzzle, spoon and fork, something that most people have a wallet full of credit cards, money and some change, I think that would give a lot of fine motor information and you don't necessarily have to carry it, you can use the patient's regular wallet or whatever, that seems to be a good one. Keys, set of keys, a zipper of some sort, buttons, or so these are irrespective of gender, irrespective of diagnoses, these could be some of the items. (E3)*

The high frequency items suggested by expert participants were using a electronic tablets (such as iPad), self care in general, eating, using a spoon or fork, dressing, hygiene, using a pen and handling coins. Technology related items are not seen in the hand function tests and yet are part of the daily repertoire of activities that people engage in. Self-care, eating and using utensils are generic tasks that the experts agreed needed to be included in a measure that targets daily activities. Dressing and hygiene is general were also reported as important elements of daily activities. The distinction between self-care and hygiene was not clear and thus the excerpts were coded by the term with which the expert referred to the task.

Hand function. The experts were able to weigh in on aspects of hand function that should be considered when designing the items.

> *I always ask, which hand are they using, what's their dominance like, what their sensation like. (E3)*

They discussed fine motor skills such as manipulation including in-hand manipulation, coordination, speed, performance efficiency, accuracy, tone, sensation including proprioception, range of motion, grip, grasp, pinch, release, and turning. They discussed gross motor skills such as reaching, lifting, bimanual function, arm swing when walking, and prior level of function. They also indicated specific problems such as learned non-use and pain in stroke, tremors and problems with rapid alternating hand movements in PD, neglect of one side, and fatigue with MS. They discussed theories that can affect assessment such as motor control, modulation of force, motor planning, and dual tasking.

Test. Expert participants were actively involved in the discussion regarding hand function testing. They cited examples of other tests that they have used in the clinic that included the Jebsen Hand Function Test, the Box and Block Test, the dynamometer, the Wolf Motor Function Test, the Unified Parkinson's Disability Rating Scale, the Comprehensive Occupational Performance Measure, and the Disabilities of the Arm Shoulder and Hand. Pediatric tests were also mentioned as one expert had experience working with children as well. The items from these measure might have influenced the perspectives of the expert participants however they were discouraged to use items directly from these measures in keeping with fair use of the commercially available tests. No items in this measure were taken from any other measures of hand function.

They also discussed the way they were currently testing patients.

> *Ya depends on why you were coming, why you were referred. I might while I am talking about what we are going to do, I might start moving your arm feeling range of motion and pain while we are chatting so I have kind of a clue, that what I am looking. Then I might do a thing like I might ask you to reach overhead, that kind of quick and dirty. (E1)*

They emphasized the fact that the measure would be useful if it could be quickly administered. The experts agreed that speed of task was important but quality of the task was equally important.

> *If it takes too long its not going to be functional. If it takes you 40 minutes to get your shirt your shirt on, you are not going to do it.*
>
> *..maybe what they are doing that you are concerned about is kind of unsafe, the way that they are leaning over to get their shirt on they are going to fall off the bed and its unsafe for them.*
>
> *But because if you saw them in 6 months they cannot do that compensation anymore well enough so now they cannot do it. (E1)*

They also noted that quality was important for progressive conditions like PD or multiple sclerosis as well. The measurement expert expressed the need for a variety of items that would tap into function (e.g. grasp) and function issues or problems with function (e.g. tone).

> *Okay, so from a measurement perspective you would want to have a variety of different tasks that we will do really tap into the different functions. (E2)*

There were other comments about giving the patients a choice of the activity they want to choose for were noted but not implemented for actual testing. However, it was noted that the individuals will get a choice of which task of the test they want to track over time that can give them information that might be meaningful to them.

The experts also expressed the need for the self-report section to come before the performance-based section.

> *I think the questions come first. I think, you have this patient. You find out why they are there, what is the problem. (E1)*

The experts also expressed a need for individualized item to track progress.

> *In my opinion I think for me it's so heavily guided by what the person tells me in their history that its difficult to generalize sometimes. I mean the general things we have talked about like reaching, grasping, in-hand manipulation, strength, sensation, those are like the basic things that I think we look for in everyone. But the rest of the finer things are so heavily guided by what they like to do you know, were they golfing before, were they fly-fishing before. So then its heavily guided by individual interests and what they are used to doing. (E4)*

Since, this could not be done in a performance-based assessment (to maintain standardized format for consistency of testing), it was decided to incorporate this in the self-report section. Moreover, since individual interests play out in higher level tasks such as work, school, and leisure, it was logical to provide an individualized item for these categories. This was built into the measure.

The experts also expressed a desire for the measure to have low equipment needs, have a standardized protocol, be able to track progress, familiar tasks, use across diagnosis, and incorporate universal instructions. The two domains of gross motor and fine motor were identified by experts and keeping a standard expectation of patient/client able to maintain sitting balance or be ambulatory needs to be set up. The experts also identified problems with current tests such as taking more time, not providing much added information, cost money, cannot find pieces of test equipment or directions that are needed, require a lot of set up time, lack gross motor skill measurement, need to guess function from scores, and the presence of too many items.

I don't know, but I would say that the big problem I see with these standardized tests for me is they take more time, and I don't get any more information than I got before. And they sometimes cost a lot of money and you cannot find the piece that you needed. You know. (E4)

Or the directions. (E3)

Or you have just an hour and I am spending 30 minutes on this test, setting it up. (E4)

Experts also agreed that the patients should be given time to understand the item, but that they may not need as much time the next time you administer the item. They agreed with giving cues or practice trials or written directions if needed. When asked about scores, there were indications that scores would be used to guide intervention but there were also strong feelings about scores needed to support documentation requirements for insurances.

I don't necessarily think practitioners use the scores, they look at what is happening and act on what is happening, it's the insurance companies who want the scores. (E2)

The other bodily functions that that affect performance on the Hand and Arm Function Measure were discussed by the group and included visual perception, visual acuity, balance, cognition, motor planning, standing, changed dominance, walking, and exercise. The measure would have to be designed to reduced variance caused by these factors on arm and hand function.

The item writing and modification phase was planned after patient/client focus groups. This was done to incorporate patient/client perspectives in the development of the construct and items before item writing.

Discussion and Conclusion

The expert panel focus group was able to provide insights into the deficiencies in existing measures, purpose of the new measure, the measurement criteria, the items that would need to go into the measure, factors that would affect hand and arm function. They were able to provide insight into conceptual understanding of hand use versus function, simulated versus real life tasks, and compensation for functional deficits that can affect measurement of hand and arm function. This expert panel study and its systematic investigation of the opinions of therapists in a qualitative research format contributes important information to the body of literature on hand and arm function related to assessment.

There is a need for development of measures that are valid in neurological conditions (Kraft et al., 2014). However, no studies have utilized the qualitative methodology to formally interview therapists and measurement experts to inform test development. Some existing measures (i.e., Disabilities of Arm Shoulder and Hand) that have utilized reviews by panels of experts (Hudak et al., 1996), have done so without explicitly documenting the results of their involvement. Some other existing measures have been developed by occupational therapists and item development has been done based on other measures or psychometric analysis (Chen et al., 2005) without involving other disciplines such as physical therapists or measurement specialists who have experience with a specific diagnostic population.

Such expert panel members are referred to as ‘subject matter experts’ or ‘content evaluation experts, by some authors in psychology (Lawshe, 1975). The experts were all women, which is representative of the make up of the field of physical and occupational therapy (American Occupational Therapy Association, 2010; Washington University in St. Louis, 2009).

There is a chance that the lack of inclusion of male experts however could create a bias in the items suggested and perspectives regarding daily activities.

Stakeholders are the individuals impacted by the use or effects of the measurement process. Since therapists are the primary individuals who will be administering this measure, involving these stakeholders early on in the construct development phase will provide evidence towards the face and content-related evidence for validity of the scores of the measure. The use and interpretation of test scores by stakeholders constitutes aspects of the construct-related evidence for validity of the scores of the measure (Messick, 1994). The expert panel input in this development phase ensures that the scores are intended to be used for the purpose outlined by the stakeholders.

Table 3.1. Construct Scheme to Represent Hand and Arm Function Dimensions and Items.

ICF	*Dimensions*	*Hand Function*	*Self-report items of Hand and Arm Function Measure*	*Performance-based items of Hand and Arm Function Measure*
Activity and Participation	Gross movements in daily activities	Transport/ reach	Toilet hygiene Washing hair Getting dressed	Hand to lower back Hand to back of head Grabbing jar at shoulder height
		Non –prehensile skilled movements (pushing, carrying)	Clap hands Shaking hands when greeting Holding an open book Lifting grocery bag	Bean can task – grasping and carrying
	Fine movements in daily activities	Prehension – Grip & pinch	Lifting gallon of milk Lift cup with liquid in it Open door lock with key Open childproof bill bottle Open & close jar Feeding yourself Cutting finger nails Items in & out of wallet Clothing fasteners Open can Brushing teeth Putting on watch or jewelry Cutting vegetables Opening milk carton	Coffee mug Open lock with key Open spice bottle Open jar Spoon use Address writing Pennies in bowl – pick & place Peas in bottle
		In-hand manipulation	Handling credit cards or money	Pennies in bowl – pick & hold
		Non –prehensile skilled movements (pointing, turning)	Open & close Ziploc® bag Turning door knob Press buttons on phone or remote control, keyboard Computer mouse	Open lock with key, open spice bottle
		Active haptic mode	Using a touch screen on a phone, tablet, computer, laptop	Sensory test - stereognosis
		Passive tactile sensing	Sensation loss Hypersensitivity	
Body Function/ Structure	Symptoms interfering with hand function in daily activities	Low functioning: Non-movement related aspects	Pain Spasticity Tremor Slow movements Deformity Stiffness Weakness Coordination Sleep Fatigue	
Participation	Occupational roles	High functioning: High demand for productive function	Work Childcare School Leisure	

Table 3.2. Design Pattern Attributes with Corresponding Messick and Toulmin Argument Components and Application to Hand and Arm Function Measure.

Design Pattern Attribute	*Messick Assessment Argument Component*	*Toulmin Assessment Argument Component*	*Application to Hand and Arm Function Measure*
Rationale	Student Model/ Claim: *What construct should be assessed?*	Warrant	Study of available measures in the systematic review, ICF framework, NIH focus on patient/client reported outcomes and NIH toolbox efforts to support similar developments in the field.
Knowledge, skills and abilities		Claim	Hand and arm function can be measured with a combination of: 1) Clinician's observation of a patient/client performing timed non-prehensile and prehensile tasks, and 2) Self-reported rating of symptoms and daily activity tasks related to hand and arm function.
Additional knowledge, skills and abilities		Claim/Alternative Explanation	1) Personal factors 2) Body structure factors 3) Cognition and other body function factors 4) Environmental factors
Potential work products	Evidence Model/Actions: *What behaviors should reveal the construct?*	Data about student (patient/client) performance	Performance section: Patient/client will participate in tasks that will be scored on a rubric to generate scores Self-report section: Patient/client will mark or indicate on the questionnaire interference of hand and arm function issues in daily tasks.
Potential observations			Performance section: Time, assistance, compensation and completion are observed by a skilled observer and recorded in the scoring rubric. Self -report section: Interference of hand and arm function issues in daily tasks and interference of symptoms on hand and arm function will indicate the functional level of a person.
Characteristic task features	Task Model/Situation: *What tasks should elicit those behaviors?*	Data about assessment context/situation	Performance section: The daily activity task features are broadly divided into gross and fine movement dimensions. The aspects of sensation, strength, grasping, and fine hand use aspects were considered based on expert and patient/client feedback. Self-report section: The task features are broadly divided into ICF categories of body structure and function, activity and participation. Within the activity and participation items related to daily tasks, the tasks were divided into self-care, kitchen, technology, miscellaneous, work and leisure.
Variable task features			There is a range of complexity of tasks that a person can engage in that will give information about low to high trait of hand and arm function is demonstrated.

Table 3.3. Comparison of Original and Revised Construct Definition, Impact of Revision on Assessment Design, and Studies to Support Evidence for Validity of Scores

Original Construct definition	*Revised Construct definition*	*Impact of revision on assessment design*	*Impact of revision on studies to support validity of scores*
Measures hand use with fine motor tasks	Measures hand and arm function – Broader outlook includes gross and fine movement domains.	Reduces construct under-representation, develop arguments with a broader definition	Includes individuals with high and low levels of the trait of hand and arm function. Extends the use arguments due to more coverage of the range of impairments.
Includes one dimension - fine motor tasks	Includes two correlated dimensions in performance-based section, i.e., fine motor and gross motor; two correlated dimensions in self-report section, i.e., body structure and function and activity and participation	Task model would include fine and gross movements, symptoms, daily tasks and become more robust, reduces construct irrelevant variance	Concurrent evidence supported by correlations between the measure and other hand function measures for both performance-based and self-report sections will need to be expanded.
Sub domains for items Control: Strength, stability, mobility, pain and coordination.	Sub-domains defined in each dimension using Jones and Lederman's framework: Transport, reach, prehensile, non-prehensile, passive sensation, active haptic mode. Low and high functioning sub-domains added.	Reduces construct under-representation, arguments need to develop for interpretation, use, value and consequences	Discriminant evidence will need to be established by selecting measures that are truly different considering that the definition of the construct is now wider.
Includes tasks (not clearly defined)	Includes daily life tasks	Impacts the use argument and consequences	Predictive evidence for outcomes predicted by this measure will need to be selected based on the daily-activity nature of the tasks.
Diagnoses not specified	Includes neurological conditions with diagnostic groups specified.	Impacts generalizability, use argument and consequences	Population selected for studies will be specific to only the specified neurological conditions. Generalization of test results to only those diagnoses is reasonable.
Response process – 5 choices with 'don't do' option	Response process – 3 choices in self-report and 5 choices in performance-based. Scoring to exclude missing items.	Improved response process Reduces construct irrelevant variance	Selection of scoring model can impact statistical analysis.
Adults with hand function problems	Adults – includes male and female, all cultures, rural and urban setting, inpatient and outpatient settings.	Improved validation of interpretations to the varied subgroups	Sample of individuals to be included in studies needs to be stratified based on gender, culture & settings.
Consequences – not developed	Consequences – Intended: Intervention outcomes, determine level of impairment, need for services. Unintended: Patient or therapist's perceptions of level of impairment or outcomes	Claims to be made about the patient/client will be impacted.	For intervention – minimal clinically important difference will need to be estimated. For level of impairment – cut off scores will need to be established. For need for services – data related to qualification for services will need to be compared.

Note: Aspects of construct definition that did not change: 1) Obtain data from performance-based and self-reported methods, 2) Use commonly available test items, so therapists can assemble their own items for the test 3) Low testing burden on therapist and client

Table 3.4. Interview Guide for Expert Panel Focus Group.

Semi-structured questions: Experts
<u>*About the construct*</u>
• What came to your mind when I said arm and hand use?
• What if I would have said arm and hand function? What would that mean to you?
• Can you tell me what is important to measure about arm and hand use? Why? How?
• Do you remember any tests of arm and hand function that you might have encountered in the past? Can you tell me about them? What were the aspects that you liked or did not like about them?
• Do you think, it would make sense to have a test of arm and hand use include both a questionnaire and some timed activities? Why or why not?
• In what form would you like to inform the results from a test of arm and hand use?
<u>*About domains of hand function*</u>
• Can you tell me what are some of important aspects or domains of hand function?
• Can you talk about how we can gather information about these domains using questions about daily tasks?
• Can you talk about how we can gather information about these domains using time performance tasks?
• What about arm and hand use in standing tasks? Does it matter?
• What about sensation?
<u>*About specific neurological conditions*</u>
• Now we will shift gears and talk about specific neurological condition and how that affects arm and hand use.
• Stroke and arm/hand use?
• Traumatic brain injury and arm/hand use?
• Multiple Sclerosis and arm/hand use?
• Parkinson Disease and arm/hand use?
<u>*About items on a test*</u>
• A test cannot have too many items. What are some of the daily activities that might be very important to measure?
• Do we need to ask about bilateral activities?
• Do we need to ask about dominance? Before or after the diagnosis?
• What about pain?
• What about fatigue?
Additional questions and discussion may follow based on the level of detail provided by each group member.

Table 3.5. Demographics of Experts

Expert	*Degree and Profession*	*Clinical Practice*	*Clinical Experience*	*Research Experience*	*Common Neurological Conditions in Clinic*	*Other related Information*
E1	MS, Occupational Therapy	Active in pediatrics	40 years	N/A	Brain tumors, brachial plexus injury, cerebral palsy, epilepsy, peripheral nerve injury, polyneuropathy, stroke, and traumatic brain injury	Teaching experience
E2	PhD, Measurement	N/A	N/A	30 years	N/A	Teaching experience Published research related to measurement in rehabilitation
E3	BS, Occupational Therapy	Active in adult care - neuro-rehabilitation out-patient	30 years	N/A	Brachial plexus injury, dementia, essential tremor, Guillian Barre Syndrome, multiple sclerosis, Parkinson Disease, stroke, and traumatic brain injury	N/A
E4	PhD, Physical Therapy	Active in skilled nursing facility	11 years	7 years	Brain tumors, dementia, multiple sclerosis, Parkinson disease, stroke, and traumatic brain injury	Teaching experience Published research related to measurement in rehabilitation

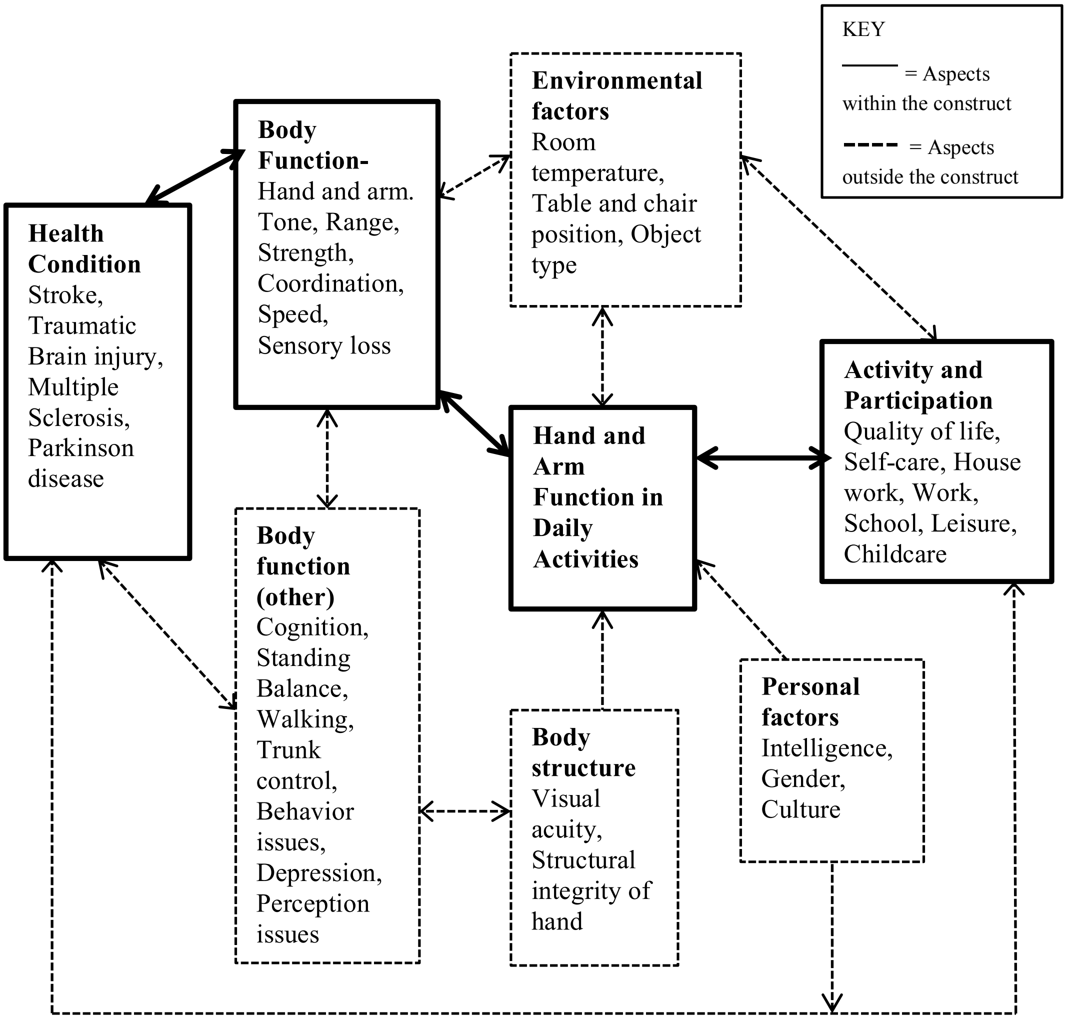


Figure 3.1. Nomological Network for Hand and Arm Function Measure

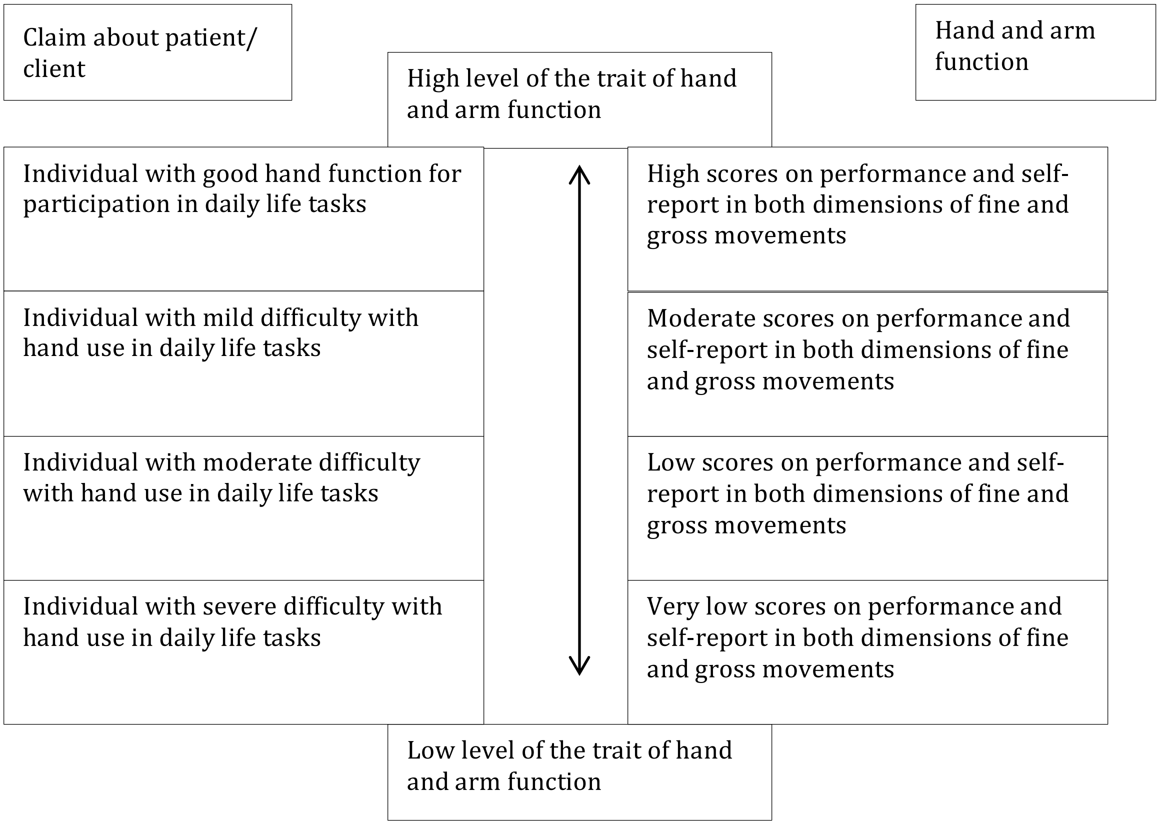


Figure 3.2. Construct Map for Hand and Arm Function Measure

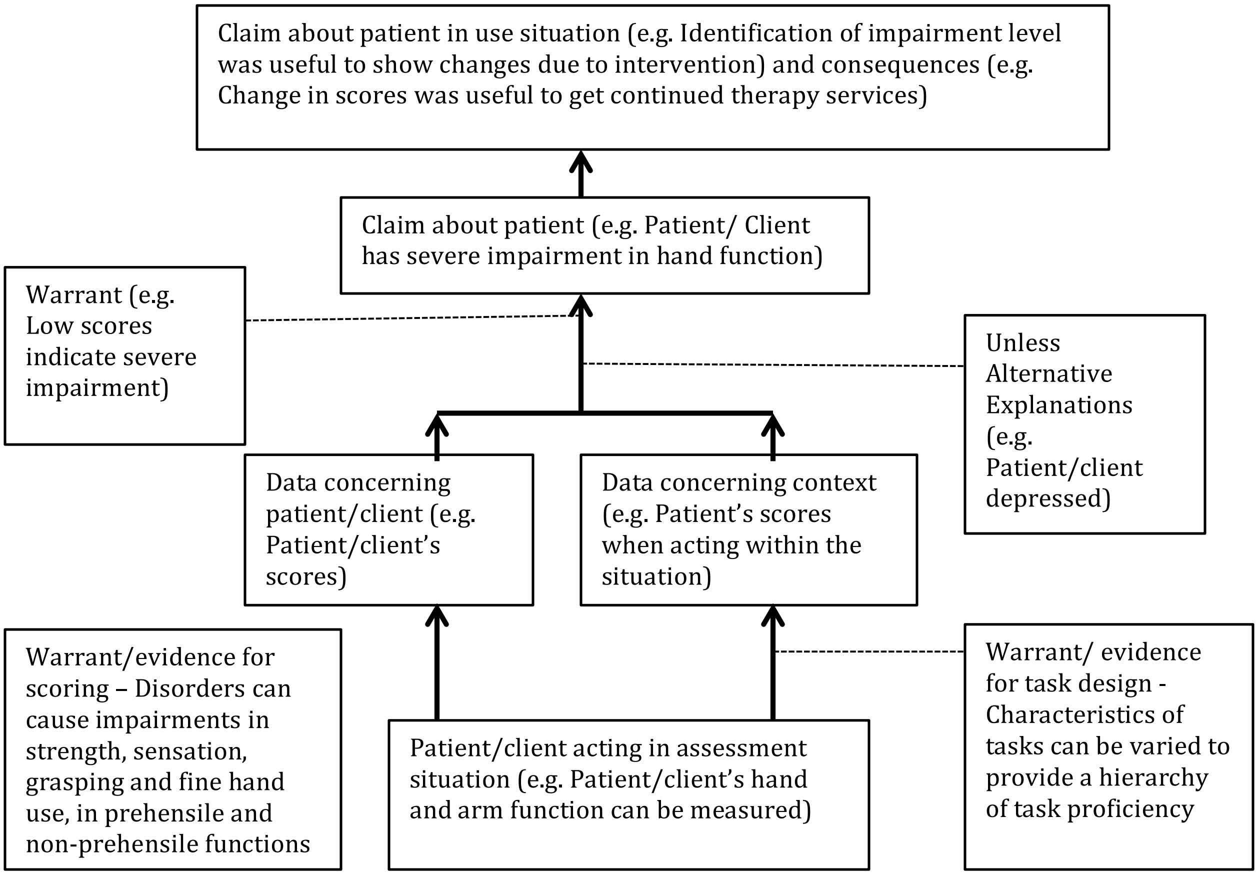


Figure 3.3. Toulmin Diagram for Hand and Arm Function Measure

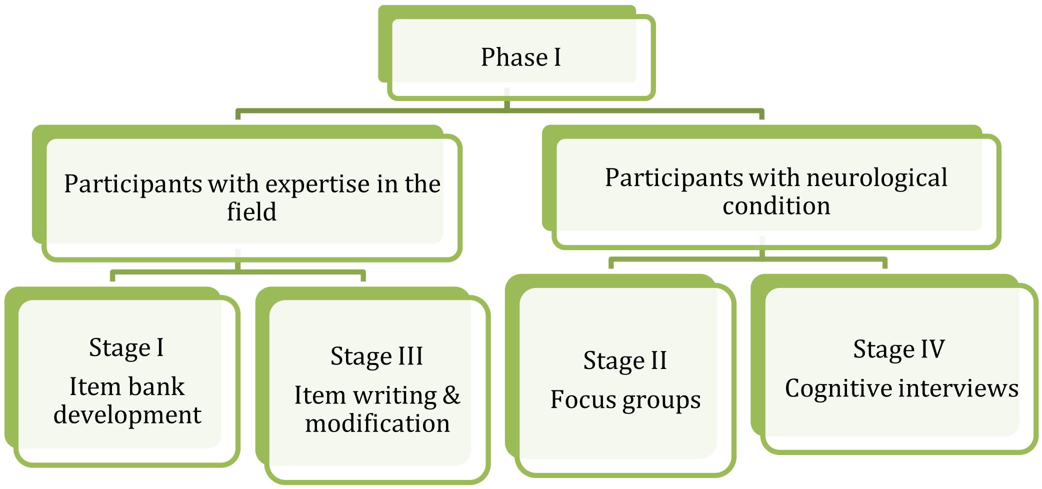


Figure 3.4. Representation of the Four Stages in Research Study

Chapter 4: Focus Groups and Cognitive Interviews Study

Abstract

Background and purpose: A new measure of hand and arm function in daily activities is being developed for use in people with neurological conditions. For the measure to accurately reflect the difficulties experienced by people with neurological conditions, a qualitative study was designed to gain their perspectives. No study currently exists that explores the experiences of people with neurological conditions in relation to their hand and arm function in daily activities.

Objective: This study aimed to explore the views of adults with neurological conditions regarding hand and arm function in daily activities and their opinion about a preliminary set of items to develop a new measure.

Method and Results: A descriptive qualitative approach involved seven focus groups and eight cognitive interviews. Twenty participants were recruited among four diagnostic conditions with purposeful sampling and included seven people with stroke, two with traumatic brain injury, six with Parkinson disease, and five with multiple sclerosis. All interviews were transcribed verbatim and thematically analyzed. The focus group discussion helped identify aspects of hand function that were relevant to the participants with neurological condition such as sensation, strength, grasping, and fine hand use. The feedback from focus group participants helped develop the item bank. The item bank from participants with neurological condition was then pooled with the items from expert panel study previously conducted to identify the high frequency items to be included in the preliminary set administered to the participants in the form of one-on-one cognitive interviews. The cognitive interview study found that self-reported and performance-based aspects of function were well received by the participants. The self-report

items needed definitions of the terms included in the questions, an additional response option and a comment box, and change in the order of presentation of items with bimanual followed by unimanual items. The performance-based items needed to capture difficulties on the affected side along with preserving the bimanual nature of the tasks.

Conclusions: The focus groups and cognitive interviews facilitated the grounding of the measure in the views of participants with neurological conditions. Thus contributing to the face and content-related evidence for validity of scores of the new measure.

Introduction

A new measure of upper extremity function in daily activities – Hand and Arm Function Measure (HAFM) - is being developed for use in people with neurological conditions. For the measure to accurately reflect the difficulties experienced by people with neurological conditions, a qualitative study was designed to seek their views. No study that explores the experiences of people with neurological conditions in relation to their hand and arm function in daily activities was found in the current literature.

This study was part of the first phase of the broad research project for development of HAFM. The first phase consisted of four stages; i.e., I) item bank development with expert panel focus group discussions, II) focus group discussions with participants with neurological condition, III) item writing and modification, and IV) cognitive interviews with participants with neurological conditions. The participants with neurological condition were involved in Stage II and IV.

Purpose

The objective of this study was to explore the views of adults with neurological conditions regarding hand and arm function in daily activities to develop an item bank and to gather their opinion about a preliminary set of items to measure hand and arm function. The broader objective of this research is to develop a new measure of hand and arm function.

Method

Qualitative descriptive research methodology of focus groups and face-to-face semi-structured interviews were used to obtain in-depth data from participants with neurological conditions. These methods were selected to facilitate a group discussion, gather ideas about potential items through group validation, identify commonly used terms for the medical condition, and identify perceptions regarding measurement of hand and arm function. The neurological conditions of interest were stroke, traumatic brain injury (TBI), multiple sclerosis (MS), and Parkinson disease (PD). These diagnostic categories were selected to represent progressive and non-progressive neurological conditions. Over a lifespan, MS and PD have a progressive course and stroke and TBI have a non-progressive course of development. These are the diagnostic categories most commonly encountered by therapists in the rehabilitation setting. Approval from the University of Washington Human Subjects Division was obtained prior to commencing the study.

Participants. The participants were recruited with flyers posted at University of Washington clinics (Seattle, WA), Evergreen Health clinics (Kirkland, WA), research participant recruitment websites, and contacting therapists in the community to inform about this study. Also research study information was posted in newsletters for registries available through various organizations such as University of Washington Traumatic Brain Injury Model System.

The purposeful sampling of participants was done with the following inclusion and exclusion criteria. The inclusion criteria were participants: 21 years of age or older, overall good health, mild, moderate or severe difficulty using one's arm or hand that interfered with daily activities, had been diagnosed by a physician to have either stroke, traumatic brain injury, Parkinson disease, or multiple sclerosis, diagnosed at least 6 months or prior, and able to arrange travel to the University of Washington for the research study and were an English language speaker. The exclusion criteria were: concurrent other neurological condition(s) that would affect arm or hand function, currently receiving physical or occupational therapy, limb loss or any injury to the arm or hand within the last 6 months, uncorrected vision or hearing problems, and memory impairment as defined by having more than two errors on the Pfeiffer short portable mental status questionnaire (SPMSQ) (Pfeiffer, 1975). The SPMSQ has 86.2% sensitivity and 99.0% specificity (Erkinjuntti, Sulkava, Wikström, & Autio, 1987) in community dwelling residents and people with dementia for the cut-off scores used to screen participants in this study without the correction for education.

The goal was to have an equal representation of each of the four diagnostic groups and gender among participants. A varied range of hand function difficulties among 16 participants in focus groups and eight participants in cognitive interviews, spread out equally among four diagnostic conditions was planned. The sample size of focus groups was intentionally kept larger to allow thematic saturation (Marshall, 1996). The sample size of eight for cognitive interviews was determined to be sufficient as the focus groups and cognitive interviews in this study were to be used to qualitatively assess individual items prior to their inclusion in the HAFM, not to perform any statistical procedures.

Procedures. Participants came to the University of Washington Department of Rehabilitation Medicine to participate in the focus groups and cognitive interviews. At the beginning of the cognitive interviews, demographic data was obtained and the following standardized self-report measures were completed either at the beginning or end of the session: MAM: Manual Ability Measure-20 (MAM) (Rallon & Chen, 2008), Patient Reported Outcome Measure Information System® (PROMIS) Fatigue short form 7a version 1.0, PROMIS Depression short form 8b version 1.0, and PROMIS Pain Interference short form 6b version 1.0 (Broderick et al., 2013). Focus group and cognitive interview data were collected from July 2014 to August 2015. The focus groups and cognitive interviews occurred approximately one year apart.

The focus groups were aimed at exploring what aspects of hand impairments affect activity and participation and what the participants' experiences have been with formal or informal testing of hand function. For example, one question in the focus group was, 'What are some of the daily activities, very important to you, that you cannot do with your arm and hand?' A semi-structured interview guide was developed for focus groups and is presented in Table 4.1. In the cognitive interviews, participants with neurological condition completed 10-15 performance tasks and self-report items, and 'talked aloud' about their perceptions about the tasks and items in an interview. The interviews at this stage were focused on item and task clarity, response choices, and context perceived by the participants with neurological condition in response to the question. A semi-structured interview guide was developed for cognitive interviews and presented in Table 4.2. For example, one question in the cognitive interview was 'How could the task instructions be improved?'

Materials. For the cognitive interviews, the materials (shown in Figure 4.1) needed for administering the performance-based section of the HAFM cost $46.56 and included: Measuring tape, paper tape, teaspoon, mug, can, pencil, pennies, lock, key, peas, jar, bottle, bowl, socks, beans, paper clip, nut, bolt, and paper pad. The details regarding the quantity, brand, size, and type of item is provided in Table 4.3.

Analysis. Data immersion, coding, category creation and thematic analysis were used to find patterns of meaning across data (Thomas, 2006; Schmidt, 2004). An inductive approach was used to derive themes by interpreting the raw data. Codes and sub-codes were identified by two researchers (NG and DK) independently. The researcher's notes during and after the session were not coded. The cognitive interview codes were tied to the items that were administered and further coded based on keeping the item or not and changes needed to implement. To organize and manage the data, the qualitative analysis software program Dedoose version 6.2 (www.dedoose.com) was used.

Results

A descriptive qualitative approach was utilized for seven focus groups and eight cognitive interviews with purposeful sampling of participants. Seven focus groups involving two to four participants and one to two researchers were conducted. Eight cognitive interviews involving one participant and one to two researchers were also conducted. Table 4.4 presents the characteristics and demographic data of the participants. Of the total of 20 participants, six participants (participant 19, 20, 22, 23, 24, and 25) participated in both a focus group and a cognitive interview. There was fair gender representation with 45% female (9 out of 20) participants. The average age of the participants was 53.81 years with age ranging from 29 to 72 years. The educational background of 95% of the participants was high school or higher. The mean level hand function of the participants on the converted score of the Manual Ability Measure (MAM) was 61.35 out of 100 (ranging from 47.8 to 100). Figure 4.2 presents the graph showing distribution of MAM scores by diagnosis. The PROMIS fatigue, pain and depression scores of participants are also reported in Table 4.4 to better understand the neurological condition of the participant and their functional level for a useful interpretation of the views expressed in the study.

All audio interviews were transcribed verbatim and thematically analyzed. For coding and thematic analysis, there was overall good agreement (focus group: 96.09% and cognitive interview: 89.35%) and the points of disagreement (focus group: 3.91% and cognitive interview: 9.13%) were resolved by discussion by the two researchers. Disagreements identified by the excerpts that were coded as 'query' or 'discuss' in the transcripts and subsequently resolved through discussion. There was consensus between both researchers on the final coding scheme as being representative of the data.

Focus group themes. The participants in the focus groups discussed the importance of remediation of hand function. The participants shared prior experiences with therapy and life experiences in general related to their hand and arm function. There was agreement among participants that there is a general lack of awareness for problems related to hand and arm function as discussed in the following piece of discussion among participants.

> *I think this is an overlooked area and that's why (in response to thanking for being a part of this study). I was explaining to my wife, the only thing you can't do without your legs is go for a walk in the park. But if you don't have hands you can't even wipe your own butt. (Participant 11, MS)*
>
> *That's very true. (Participant 13, MS)*
>
> *Why is the wheelchair the biggest fear? There is something psychological about it. (Participant 11, MS)*

Eight key themes emerged from analysis of the focus group transcripts.

Conditions. A participant with TBI talked about his bilateral involvement, shakiness of hands, co-morbid seizures, anger issues, and his frustration with hand problems at 29 years of age.

> *I like to watch sports stuff like that. I want to get into fishing again. But I only have one arm so everyone is like, no you have two arms. But I am like no I just have one arm. Because I pretty much consider this one dead. Mainly in my head I wish this TBI, this hand would have gotten cut off or something. (Participant 18, TBI)*
>
> *I cant use my this (left) hand. I can still use my right hand to do everything. Even though I can't really use it (right hand) just because it shakes so much you know. But it's the only one I have so I use it. (Participant 18, TBI)*

Participants with PD talked about the differences among people with PD and that no two individuals with PD are alike.

The things (about differences between people with PD previously said) that Participant 7 said so eloquently, it's different for everybody. And your progression is different for everybody. (Participant 5, PD)

That's why they call it a snowflake disease. (Participant 7, PD)

Ya. (Participant 5, PD)

Snowflake disease? (Researcher, NG)

No two snowflakes are the same, no two Parkies (referring to people with PD) are the same. I don't know if it matters but its true. (Participant 7, PD)

Some of the symptoms discussed by participants with PD were ataxia, cramping, dyskinesia, micrographia, rigidity, feeling numb, freezing, slow movements, sudden jerky trigger movements, tremors, voice changes, fixed postures of the hand, and issues related to progression of the disease.

I keep trying to do this, (showing a hook fist) to not show my tremors. (Participant 6, PD)

Oh the 'Parki-claw.' (Participant 7, PD)

Participants with stroke talked about the tone, co-morbid seizures, the road to recovery of function, the change of dominance, and coping strategies used by them. Participants seemed to appreciate sharing the stories of their recovery in the group. Participants with MS talked about the episodes, relapses, heat intolerance, optic nerve problems, variable presentation and issues with progression. These condition specific symptoms prompted the introduction of a symptom inventory at the beginning of the self-reported section of the HAFM. Further, to acknowledge the different symptoms on the two sides of the body, the questions would separately ask for left and right hand symptoms for all conditions.

Compensation and adaptation. There was a lot of discussion about the compensatory strategies used by participants, including demonstrations of adapted tasks such as picking up golf tees, playing with a ball, and holding two grocery bags. As a result, compensation was included as one of the aspects within the response options and therapist was provided space to record notes during session related to compensation or abnormal movements.

Dominance. Some participants reported the strong sense of loss due to loss of their dominant hand function.

It is so strange. I want my left hand back so bad…(Participant 8, stroke)

Some others reported that their non-dominant hand function was not well trained since the dominant hand overpowered them. It was clear that function was very different for people with dominant hand more affected than the non-dominant hand.

Items. There were 160 item codes that generated 145 items in the item bank presented in Appendix D. Some codes were grouped together to form items, for example, handling bags, purses and backpack were grouped together as a subcategory within 'object handling.' There were nine main subthemes within which the items were grouped: i.e., carrying, communication, instrumental activities of daily living (IADL), movements, object handling, reach, recreation, self care, and transportation. Carrying involved being able to carry items like groceries, coffee cup, wine glass and bag from one place to another and involved endurance of upper extremity to hold a sustained contraction and keeping the hand steady during movements involving rest of the body. Communication represented using hands as tools of communication for clapping, shaking hands with another person, and writing.

My handwriting has definitely gone downhill. I can't read it any more. I have a diary, but I can't read what I have written on the diary. (Participant 14, PD)

But I would say clapping looks a little sloppy and it's been that way for a few years. I noticed it because when teaching children and clapping along, I kind of have to find a flat part to clap. (Participant 17, MS)

My wife laughs at me when I am clapping, she is like - "Is that it" (making a low and flat sounding clapping sound). At the Sounders game, I am like cheering, that's pathetic (claps). (Participant 16, MS)

IADL tasks include activities related to independent living such as childcare tasks, grocery, housekeeping, kitchen tasks, laundry, work and others. The group of individuals in this study were community dwelling individuals who wanted to regain or maintain their ability to engage in IADL tasks. Movement tasks involved the movements that were articulated by the participants such as holding, lifting, opening and closing hand, picking up, placing on table, and tapping on table. Object handling included all the different objects that were mentioned by participants as challenging to manipulate due to hand and arm issues. Some examples of objects were a ball, bottle, cans, child safe medication bottle, coins, credit cards, cup, doors, jars, and keys. A sub-theme within object handling was handling of technology related objects such as computer, tablet, phone, keypad, mouse, typing and touch pad. Reach included reaching to a shelf and taking a box off a shelf. Recreation included various leisure time tasks that were relevant for participants, and were very varied. If the measure was to address recreation, participants expressed there had to be a way to capture the specific leisure task that was relevant to the individual. This was also evident for the theme of work as well. Participants expressed offering a way to capture the work related tasks that are most important to the participant emerged as an important component of the measure.

Hand function. The participants had varied perspectives on the various functional aspects of the hand.

Bilateral tasks. Participants were trying to use the affected hand in functional tasks and reported that they resort to doing one-handed tasks only when faced with time or activity constraints.

> *Ya, I was always right handed. And I said the main thing that was difficult was cutting. Doing laundry, I have a hard time pulling it out of the dryer. I could use both hands but this hand is very weak so if I need to get things done fast I just use my dominant hand. Which is not helping my non-dominant hand. (Participant 10 with stroke)*
>
> *You can't clap. (Demonstrates clapping) (Participant 2, Stroke).*
>
> *So does that bilateral function matter? Two handed. (Researcher, NG)*
>
> *Oh yes, yes. It's a hassle to do one handed. (Participant 2, Stroke)*

Doing activities with both hands was important to the participants and to some extent it was their expectation from a therapy session that the therapist would want to learn more about their affected hand and aim to tailor the session to improve function.

Low level of hand function. The participants expressed their perspectives about very low hand function in terms of relief of symptoms and passive movements.

> *I have, in trying to make my left arm do things, it activates spasticity, which I get Botox…I actually don't have any hand function except, I can push something. …The Botox relaxes the muscles? It makes it so that even if I cannot do it willfully, I can do that*

(shows passive opening the hand and straightening the elbow, flexion of shoulder). It doesn't hurt. (Participant 2, Stroke)

The passive movements were not included to avoid ceiling effects in the items in the measure.

Strength. Hand strength was one of the most commonly identified aspects of hand function and included grip and pinch strength in the discussion. The most commonly tested aspect of hand function in the clinic that all the participants could identify was grip testing. Some hand function measures such as the Wolf Motor Function Test include grip testing with dynamometer along with the other items in the measure. However, it was not considered for inclusion in this new measure of hand and arm function because of its instrumented nature. Daily activity tasks like picking up bean can and using a key were considered representative of grip and pinch strength respectively for this measure.

Sensation. The participants linked aspects of hand sensation to the hand function discussion. Aspects such as estimating the force for gripping, proprioception for knowing the hand position, tingling, and having to use vision to compensate were discussed.

I have to really look at it (the paper cup). Keep my eye on it. But if I turned away, I wouldn't realize that I have just crushed it and spilled it. But if I really focus I could do it (keep it) right there and put it there. But I wouldn't get up and walk with it or anything. (Participant 16, MS)

The final items thus included sensation aspects to get a patient and therapist reported measure of sensation.

Day to day changes. The participants brought forth this unique perspective of day-to-day changes in their hand function and even general health. This view was shared by participants with MS, PD and stroke, as seen the following quotes.

> *So there isn't a single activity yet that is completely gone. At times I have more or less difficulty. Or more or less sensation and I am just constantly, you know stretching and very conscious of my hands. (Participant 11, MS)*
>
> *So I mean good days and bad days, we all have them regardless of what kind. (Participant 15, PD)*
>
> *So its hard to take that variable out of the equation. (Participant 14, PD)*
>
> *Well some days I wake up and you just cant sit still and some days I can be fairly calm most of the day and it affects your energy level because this takes energy when you are moving all the time. (Participant 14, PD)*
>
> *Ya sometimes it feels like it gets stronger but I feel I do have sensation in it but I can feel right now. Some other times I am not aware where it is. Like my boyfriend was saying he will say when we are sleeping I hit him in the face. (Participant 10, Stroke)*

The self-report questionnaire seemed sufficient to capture the variations in function. Thus, this feedback was not incorporated for the measure during cognitive interviews. As our understanding of the way to incorporate this feedback developed after the cognitive interviews and was supported by participant 26 with TBI's suggestion of asking in the questionnaire if this was a typical day, this question was incorporated in the revision after cognitive interviews.

Functional categories. Participants also identified aspects of hand function such as strength, dexterity, accuracy, coordination, endurance, fine motor, hand movements, range of motion, reaching, speed, tone, tremor, turning the hand, pain, and neglect.

So you have strength and dexterity, and accuracy. (Participant 5, PD)

What's accuracy? (Researcher, NG)

Being able to punch the button as opposed to the one next to it. And then repetition. (Participant 5, PD)

Participants helped describe these aspects. Patient's familiarity with some of the commonly used therapy jargon also supported the idea of including some therapy jargon within the measure. For example, people with PD were familiar with tremors and incoordination and people with stroke were familiar with tone and spasticity.

In the discussion of sitting versus standing for doing tasks, some participants' thought standing and doing tasks was harder due to poor balance.

Oh ya you've got your balance to worry about to stand. (Participant 4, PD)

Standing ya and now you are dual tasking. (Participant 7, PD)

Another participant thought sitting and doing tasks was harder since her shoulder would work best in standing to do the task and be able to generate more force.

Depends on a task I guess. If its sitting at the kitchen counter, you can stand and do it, that might be similar but if its coming from your arm the function is going to be different here than up here… Say like when you are cutting, your shoulders would be working all the time sitting. (Participant 1, Stroke)

I was thinking its easier standing. (Participant 2, Stroke)

For the measure, all tasks were administered in sitting to avoid the influence of standing balance on the items. Sitting balance was not discussed in this group as none of the participants had trunk control issues. Sitting balance might need more investigation through qualitative data in the future for this measure.

Hand use versus function. One of the questions from semi-structured interview guide was related to a discussion of terms "hand use" versus "hand function."

I don't know what the difference is. (Participant 2, Stroke)

Function to me says it can function. (Participant 1, Stroke)

Function is specific, like you have fine motor dexterity, more than usage how much should I be able to use it. What does use mean? Because people compensate and are able to get the task done but the motion so that they are not using the certain muscle. (Participant 17, MS)

I like the word function, there is a much greater meaning than use. Use is like I think is so general. (Participant 16, MS)

There was a repeated affirmation from participants that they understood what a hand function measure would entail. Hand use seemed more ambiguous to the participants.

Medication. Medication management seemed to be a very important piece of discussion for the people with MS and PD. Moreover, for PD there were direct effects of medication wearing off referred to as the 'off' phase versus the 'on' phase of good function under the influence of medications. It was clear from repeated reporting of medications and their effect on function that it was important to find out if the participant was in their 'on' or 'off' phase. For this, it was important to know more about the medication taken before the administration of a measure. Thus, space to report medications was added to the patient information section.

Test. We were interested to find out if participants had any opinions about their hand function being tested. Some voiced their concern about testing.

I am not sure I capture how uncomfortable it can get to be measured…We don't need the measurement. It is for the clinician, not for us. It is not even for the clinician, it's for the clinician's supervisor to tabulate results. (Participant 11, MS)

Don't test, don't just test without any offering. (Participant 11, MS)

The participants indicated they wanted the outcomes of tests being used in some form when probed about results of the tests that participants have undergone.

Well I don't think I have seen any results of any of the tests (of hand function) I have done. I mean there is not a pass/fail. That you just plug the information in and somebody grades it in some aspect but I don't think it relates to me as an individual. (Participant 15, PD)

Would you like to see those? (Researcher, NG)

Doesn't make a difference. (Participant 15, PD)

Some participants did not want to know the results possibly because of a progressive disease course. Thus, it was decided that the participant should be offered the choice to know about their test results. Some participants also requested a comparison of the current performance to past performance and a comparison of performance between the two upper limbs.

The normal range of scores and where I am at now. Ya do it separate, one arm and then the other arm. (Participant 18, TBI)

Participants offered many suggestions about the measure. They suggested, it would be helpful if the self-report section be administered before the performance-based section; and that pictures be used to augment verbal instructions. The participants also spoke to the importance of having items that are relevant to their lives and that a variety of tasks be included. They further expressed the importance of being able to track the performance of their left and right hand separately and be able to track change over time.

When asked about being timed, some of the participants mentioned that they did not like being timed at all since it adversely affected their performance as discussed in the following quote.

Well when I am timed, I find myself not able to do it as well. (Participant 14, PD)

The stress of the time, you have to get done in a certain amount of time and then all of a sudden I cant do it. If I have the time to do it. (Participant 14, PD)

Other participants enjoyed the competitive aspect of being timed and would try to beat it. In view of the importance of time to track small changes in performance, particularly tasks needing accuracy, it was decided that only dexterity related tasks would be timed.

The complete item bank is presented in Appendix D. These items were then organized by the frequency with which they were coded in the transcripts. Since there was only one individual with TBI, the frequencies for the item codes were adjusted to a value of one to avoid over-representation of any one single code (e.g., the code 'Carrying a coffee cup' occurred three times for the individual with TBI, however it was adjusted to a value of one). The items from expert panel study and focus groups were pooled. In order to pool the items, the item codes were kept consistent across the two studies. The top 1/3rd of high frequency of the pooled items were selected as the final item set. The 59 identified items were then sorted into self-report and performance-based categories based on testing performance for items that are critical for movement quality, dexterity and speed. The item development process is presented in Figure 4.3.

The final set of items for the cognitive interviews had 11 symptom related questions, 34 ADL questions, and 14 performance-based tasks. The reading ease according to Flesh-Kincaid in

MS Word was 74.4% at a 5.5 grade level for version A of the self-reported section of the measure and 73.4% at a 5.7 grade level for version B of the self-reported section of the measure. The reading ease for the performance-based section was not calculated since the target audience reading the document were therapists who were familiar with higher reading levels and medical terms needed for explanations. These measures are presented in Appendix E for self-report section of the measure and Appendix F for performance-based section of the measure.

Cognitive interview themes. The participants made varied suggestions to changes to the items, wording and tasks. Most questions were presented after participants completed four to five items. The feedback from the interviews resulted in substantial changes to 39 out of 59 items with three major themes; new tasks, changes suggested for items, and changes suggested but not implemented for the items.

New tasks. Participants suggested new tasks as they remembered other tasks related to the items. These are presented in Table 4.5.

Now you don't have tying a tie on there. ... that's a big one. It's a big endeavor, (to tie a tie) it takes half an hour. (Participant 23, Stroke)

This quote highlights the suggestions for specific tasks that was important for this male participant and could not be incorporated in the measure since female participants would not be able to attempt to task. It was discussed by participants that medication sorting task into pill box for the week was a meaningful task and this being a universal task was added in the list. ICF domains of throwing catching, pushing, pulling, were not included in the measure and will be incorporated in the self-report section. Release of the fingers to release an object was not included as a specific item but the response options do take into account the ability to release the object.

Changes suggested for items. The final set of 59 items administered during the cognitive interviews had 11 symptom related items, 34 ADL items, and 14 performance-based task items. Each item was administered to five to seven participants, ensuring all items were tried by at least one individual in each of the four diagnostic groups. The items were split into version A (Self-report: 34 items, Performance-based: 10 items) and version B (Self-report: 29 items, Performance-based: 10 items) with overlapping items. The version A for both sections was administered before version B for five participants, and in the opposite order for three participants. 14 performance-based tasks were split into 21 items since some items were administered on both sides, right and left (e.g. touching the back of the head on right and left). The changes suggested by participants were also coded for the item for which the change was suggested. Participants were probed about the burden of the items and of being in a testing situation as a whole. None of the participants felt that the items were overwhelming or that too many items were presented to them.

There were six sub-themes within the implemented changes; i.e., items, response options, scoring, demographics, symptoms, and medications. The sub-themes represent the items in the self-reported section and are followed by the performance-based section to organize the information.

Items. In the self-reported section, optional items were liked by the participants but they suggested the format needed revision. Moreover, the childcare and school item did not seem to of particular interest, even among people who were engaged in daily tasks related to them. Thus, it was decided to remove the two items. The work and recreational items were preserved in a changed format that allowed the participant to list an item and rate it without rating the item in general. The item related to tying shoes was preserved and putting on shoes was removed due to confusing wording and redundancy. The item related to ‘wallet or purse’ was changed to ‘wallet or purse or pocket.’

In the performance-based section, the ‘hand to lower back’ item instruction wording was revised for reaching to the back at the waist to avoid confusion as to where on the back to reach. Pick up the coffee mug needed more clarification of ‘using the handle’ to avoid a person from not using the handle since handle needs more coordination. In the item ‘pennies on table,’ pennies to be stacked on the table or not was not clear. For the sensory item, the pictures of the objects were provided for patients with difficulty naming objects. One participant observed that the names of those items should come right under the pictures. Moreover, if pictures were shown to participant when testing one hand, for example right, the same object should not be used for the left hand otherwise the participant gets a hint as to what items might be in the sock during testing of the left hand.

The order of the sensory items needed attention. The cotton ball was a difficult item to identify for the participants and was only presented on one side occasionally. It was perceived as unfair if the affected side encountered the cotton ball. Thus, the varied texture and temperature of objects was perceived as inconsistent in the test. A clear solution does not exist to this issue; however keeping the same texture for all objects is possible. For example, if all objects were metal like key, coin, nut, or bolt then the test would be consistent for the material and remembering the type of object could be avoided with repeated testing. Some of the items in the test were bimanual items (e.g., opening a lock with a key), and participants were given a choice to use the key in any hand. Not surprisingly, participants held the key in the less affected or unaffected hand. The task was originally designed this way since the participant has a choice for which hand to use in real life. Thus, the measure would reflect their real performance. However, participants voiced their concern with this idea. They argued that their purpose in therapy is to gain more functionality of the affected hand. This can only be achieved if the therapist identifies the problem when testing them with the HAFM. Thus, the suggestion by participants to test their ability to do the task both ways, i.e. key held in affected and unaffected hand alternately, was accepted. The bean can task needed major revisions including improving the clarity of the picture provided for set up, extending the reach with the bean can beyond the two 6 inch markers, and planning pick up and place of the can to coincide with start and end of the task. Moreover, the bean can task did not intend to represent carrying and strength of grasp domains of hand function as originally planned. It did represent the carrying domain, but the grasp was redundant with the grasp for the jar, and a separate strength task such as 'pouring water from a pitcher' or 'picking up a half gallon of milk' would be added. The opening jar and opening spice bottle task involved similar movements and were redundant. Thus, the more commonly performed opening jar task was preserved and opening spice bottle task was removed.

Response options. In the self-report section, the participants suggested adding a comment section in order to indicate specific details about the tasks as they mark the options, and remember the detail at that moment. The three options for response given in the self-report section of the measure were not at all, a little and quite a lot. More than one participant indicated that their function was more impaired than quite a lot. Another category of response option 'unable to do' was added to address this issue. If response options are not clear, therapist can help the participant problem solve the degree of difficulty and this instruction needs to be added to the manual.

Scoring. In the self-report section, the numbers that were meant for therapist to calculate the score were very confusing for participants and some tried to count the score themselves. The self-report measure did not specify if the participant had to mark the response with a check or a cross. The commonly used check mark option was added.

For the performance-based section, there was no way of recording follow up assessments unless a new form was used. The assessment form needs to have the ability to record an initial, follow up and discharge assessment. Participants reported that a numerical score with nothing to compare against was not helpful. A qualitative summary that they did well on a skill such as reaching and average on a skill such as fine motor would help them start a conversation with their therapist about their goals. This was possible to implement for the performance-based section since the task design could be tailored to the function to be assessed. The combination of hand function aspects on the performance-based section and daily activity domains on the self-report section might give the clinician a better picture to further investigate in the therapy session. Moreover, it was decided that the manual would include suggestions for the therapist to offer the client/patient a chance to choose one item that is most relevant to them to follow up over time in a numerical form.

Set up. The set up needs were only for the performance-based section of the HAFM. One of the participants had a skin condition and this brought to attention the need to add explicit wording about cleaning the test items in the manual. The administration of a new measure of hand and arm function required the reading out of standard instructions to the participant, however it was necessary to tell the participant that the therapist will be reading out the standard set of instructions in order to clarify the administration process to participant.

Demographics. The demographics information was entered by the participants for the self-report section of the HAFM. A participant remarked that the 'date' be changed to 'today's date.' The word 'hand preference' was preferred over the word 'dominance' by participants. The word diagnosis was difficult to understand for some participants and will be defined in the measure.

Symptoms. In the self-report section of HAFM, meanings of the words spasticity, hypersensitivity, and deformity were not clear. However, there were some participants who were more familiar with these terms possibly because they experienced the symptom and frequently used it with clinicians. Thus, there was a need to accurately pose the question related to the specific symptom but there also was a need to define the term for people who were unfamiliar with it. It was thus decided to add descriptions of the terms spasticity, hypersensitivity and deformity in the revision of the measure. Participants remarked that on the self-reported section, the questions start out with symptoms on right and left hand and then move on to items that are assumed to be bimanual. This order was confusing for the participants. One participant suggested that the right and left hand separation should happen in the end to avoid confusion. Thus, the order of the tasks was changed to move symptoms towards the end of the measure. More than one participant suggested that a reminder that all questions are targeted towards the past week is needed at the beginning of each page. For the Ziploc® item, opening a Ziploc® was reported to be harder than closing thus closing the Ziploc® was removed from the item. It was suggested that we use the word typing as opposed to keyboarding to keep the term more general.

Medication. For self-report section, it was clear from patient reporting that not only providing space to enter the medications taken by the patient was meaningful, but also the timing and dosage of medication influenced arm and hand function and was important to note. It was decided to expand the space needed for entering medications on the patient information section before the administration of the measure. Also, patients were occasionally confused about the medication or dosage and an instruction in the manual was to be added to request patient to bring a list of medications to the therapist before session.

Changes suggested but not implemented for the items. In the self-report section, the participants suggested that symptoms should only be listed as right and left, however in its current form, the terms 'right hand' and 'left hand' remind the participant that the symptoms probed are related to hand function.

In the performance-based section, it was suggested by one participant that writing a check was a good activity for high functioning individuals with PD to bring out the micrographia. However, since the writing task is timed, it was assumed that the time factor would capture the differences between high and moderate function. One participant suggested stacking the coins to pile them up as a modification to the penny on table task. Given the potential increased length of time it would take an individual with moderate to low hand function to complete making stacks of coins, this suggestion was not implemented. One participant suggested that the items that were difficult for her should be completed with assistance to help her learn the skill. This suggestion was also not incorporated as it is contrary to the purpose of the measure to identify the limitations in function and it is not desirable to teach the tasks of the measure to the participant. Activity limitations are more appropriately be the target of therapy. For the same reason, another suggestion that was not implemented was the use of a non-skid pad or mat.

Finally, the 13 self-report items that worked well were in the subsections of self-care, technology, kitchen, and miscellaneous tasks. The current form of the performance-based items worked well for seven tasks, i.e., touching back of the head, getting the coffee mug, getting the jar, opening the jar, picking up spoon, and holding the pennies.

Discussion

This qualitative study aimed to explore the opinions of adults with neurological conditions regarding hand and arm function in daily activities and their perceptions about a preliminary set of items to measure hand and arm function. The focus group discussion helped identify aspects of hand function that were relevant to the participants with neurological condition such as sensation, strength, grasping, and fine hand use. The feedback from participants helped develop the item bank. The item bank from participants with neurological condition was then pooled with the items from expert panel to identify the high frequency items were included in the preliminary set administered to the participants during the cognitive interviews. Based on the cognitive interviews it was found that self-report and performance-based aspects of function were well received by the participants. The performance-based items needed to be tailored towards impairments on the affected side along with preserving the bimanual nature of the tasks. The self-report items needed definitions of the terms included in the questions, an additional response option and a comment box, and a change in the order of presentation of items with bimanual followed by unimanual items. Focus group and cognitive interview probing also helped shape the construct with more attention to the affected side and to the aspects of the lower end of hand function. The problems with the items resulted in modification of items or removal of the items that were redundant (e.g., opening jar and opening spice bottle) or did not work as intended by the researchers (e.g., interference of hand function with childcare).

The participants expressed liking the ‘not applicable’ option in the self-reported measure. There have been prior studies (Morgan, Amtmann, Abrahamson, Kajlich, & Hafner, 2014) where participants have reported the usefulness of having a ‘not applicable’ option. A not applicable option however, can present potential scoring issues that may reduce the quality of data (Krosnick et al., 2002). However, in the absence of a ‘not applicable’ option, participants may provide inaccurate information or leave a question blank. In both of these situations, the quality of the data may be threatened. Providing a ‘not applicable’ response for an item of hand and arm function that a participant may never encounter in their daily life (e.g., an item like ‘washing hair’ when a participant did not have hair) is consistent with the assessment framework of this measure.

This study contributes to the literature and demonstrates the involvement of stakeholders in the development of a measure of hand and arm function. This is consistent with the guidance for researchers by PCORI (Patient-Centered Outcomes Research Institute, 2012)and large regulatory bodies like the Federal Drug Administration (FDA) (U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research et al., 2006). It also aligns with the 8-stage framework presented by Velozo and colleagues (2012) who also advocated the use of mixed methods in preliminary research (Velozo et al., 2012). The use of focus groups and cognitive interviews with people with neurological conditions is the strength of the process of instrument development for the HAFM.

There were some differences among people with different neurological conditions. The people with stroke, stressed recovery and the effect of spasticity issues. The people with TBI talked about ataxia and participation in higher level work related issues. The people with PD were very concerned about the effects of medications on function, particularly tremor and incoordination. The people with MS were concerned with issues related to progression of the disease and the lack of awareness among clinicians for hand function problems. These differences have been acknowledged by some authors (Lundy-Ekman, 2012), but not used for measure development in hand function.

Some studies have used formal reading frameworks for the calculating reading level (Devraj & Wallace, 2013). The ease of reading the instructions and questions was calculated for the self-report measure, but a reading level using a standardized reading framework was not used in this study. The need to further evaluate was not sensed given the documents were at the reading grade level according to Microsoft Word's Flesch Kinkaid Reading Level of 5.5 to 5.7 for the two versions of self-report. The diverse literacy levels of cognitive interview participants based on self-reported educational level from less than high school to doctoral degree was a beneficial aspect in this study to identify the ease of language and minor nuances in wording and format.

Limitations. Not uncommon to qualitative studies, this study is limited by influences of personal researcher biases (Anderson, 2010). A detailed log of researcher's notes post session were kept with the intention of recording any biases and involvement of the second researcher during analysis were attempts made to neutralize this effect. Moreover, the primary investigator was an occupational therapist and second researcher was a physical therapist, which helped reduce some of the inherent biases related to perceptions of daily activity limitations seen in occupational therapy practice. Also, a group of 20 participants may not be representative of all the views within the community of people with neurological conditions.

There are other ways to established content-related evidence for validity of scores (Lawshe, 1975), like quantitative analyses. Lawshe (1975) has suggested creating a survey for the expert to comment on each item with one of the three responses – essential, useful but not essential, and not necessary. The responses of the experts are pooled and the number indicating ‘essential’ marked for each item is computed. Although, this approach provides quantitative data for item selection, the sample size of experts in this study was not sufficient to design a quantitative study. Moreover, no added benefit was perceived in structuring in a quantitative format over qualitative format for data gathering when the open ended questionnaire had more room for experts to present their opinions, suggest new changes and suggest new items if needed.

Some theorists (Netemeyer, Bearden, & Sharma, 2003) have argued that theoretical framework provides the best conceptualization and discounted the benefit of participant feedback. However, participant feedback supports the evidence for validity of scores and use of the measure. The field of clinical testing is very much client-centered and must consider the input of the patient or client in scale development.

The self-reported and performance-based formats of tests themselves present with strengths and weaknesses. The major drawback of self-report format is the lack of faith of clinicians in accurate reporting by the client or proxy. However, NIH initiative on PROMIS has clearly emphasized the need for more patient-reported outcomes. The main reason for this initiative is the involvement of the client in the rehabilitation processes of assessment, setting goals, treatment and discharge. The self-report itself can start a dialogue regarding the problems encountered by patient or client that can be addressed in therapy. The performance-based format is traditionally well accepted by therapists and gives the therapist an objective measure of patient or client’s hand function in the areas tested. However, it is a measure of capacity and does address the varied daily tasks encountered by the patient or client in routine activities.

The hand function level of participants was 47.8 to 100 converted score on the MAM and it is speculated that individuals with low functioning were not adequately represented in this group. Thus, it is possible that a complete assessment of function was not done. Ongoing research will need to focus on recruiting participants with lower hand function to address this issue. Also, individuals with TBI were not well represented in this group with a total of two participants. Further efforts need to be made in order to recruit more participants with TBI in future studies. Also, the regional nature of sampling may place limits on the generalizability of the findings.

Although the participants did not report that the measure was burdensome, it may not be an accurate representation of how the measure would be perceived in the clinic. The participants in this study were volunteers who seemed to care about the hand and arm function issues in neurological conditions, which was the reason they were participating in this study. Further, the participants had two hours to complete the items and discuss them, which is very different from pressures of a busy clinic. Thus, clinical testing needs to be done in the future to identify the optimal number of questions that will not feel burdensome to the participant and therapist in a clinical environment.

Conclusion

The focus groups and cognitive interviews facilitated the grounding of the HAFM in the views of participants with neurological conditions thus helping to establish the face and content-related evidence for validity of the scores of the new measure. Psychometric studies are needed in the future for measuring evidence for the reliability and validity of scores of this measure in assessing upper extremity function in people with neurological conditions for its intended use of tracking progress over time.

Table 4.1. Semi-Structured Focus Group Guide for Participants with Neurological Condition

General measurement questions
What came to your mind when I said arm and hand use? What if I would have said arm and hand function? What would that mean to you? What is important to measure about arm and hand use? Why? How? Do you remember any tests of arm and hand function that you have taken in the past? Can you tell me about them? What if a test had a questionnaire to fill out and a portion where there were timed tasks that you would have to do, would it make sense to have both in a test? In what form would you like to know the results from a test of arm and hand use?
Specific questions related to the neurological condition
How does your neurological condition affect upper arm use? How does your neurological condition affect lower arm use? What are some of the daily activities very important to you that you cannot do with your arm and hand? Do bilateral activities matter to you? What about dominance? Were you right or left-handed? Did that change with the diagnosis of your neurological condition?
Additional questions and discussion may follow based on the level of detail provided by each group member and flow of the conversations.

Table 4.2. Semi-Structured Cognitive Interview Guide for Participants with Neurological Condition

Performance-based measure questions
What were your first impressions about the tasks?
Tell me how difficult you thought the tasks were?
How could the task instructions be improved?
Tell me how relevant the tasks are to you?
What do think about having some of the tasks timed?
What if any tasks should be included in this test?
Self-report measure questions
What were your first impressions about the questionnaire?
How was the wording of the questions?
Tell me how relevant the questions are to you?
What about the time frame of "in the past week"?
Tell me how the response options worked for you?
What if anything specific to your condition should have been on the questionnaire?
Additional questions and discussion may follow based on the level of detail provided by each group member and flow of the conversations.

Table 4.3. List of Materials for Performance-Based Section of Hand and Arm Function Measure

Material	*Type*
Adult tube pair	Black adult tube sock pair Hanes® Brand: Length: 12 inches
Bean can (unopened)	Bush's Best® brand dark red kidney beans can, paper removed during test: Capacity = 16 oz., Weight = 1 lbs.
Blackeyed peas (15)	Goya® brand dry black eye peas
Bowl	White porcelain cereal bowl Correlle Livingware® brand: Capacity = 18 Oz, Inner rim Diameter = 6 inches
Coffee mug	Porcelain coffee mug with one handle on the side, Correlle Livingware® brand: Capacity = 12 oz., Height 4 inches
Coins - pennies (16)	US currency penny = 1 cent
Dry kidney bean	Goya® brand dry red kidney beans
Jar with a screw top lid (contents sealed)	Skippy's® brand peanut butter jar: Weight = 1 lbs., Capacity = 15 oz.
Key	Metal key part of the Masterlock® brand no. 2 set: Length = 1.6 inches
Key and a padlock	Masterlock® brand No. 2 padlock: Width = 1 1/8 inches, 29 mm, Shackle height = 1/2 inch, 13 mm. Key: Length = 1.6 inches
Measure tape	Stanley® brand key chain tape measure
Nut and bolt	Nut: Diameter = 1/2 inch. Bolt: Diameter = 3/8 inch.
Paper clip	Paper clip: Length = 1.1 inch or 28 mm
Paper tape	Curad® brand gentle paper tape: Width = 1 inch
Pencil (at least 6 inches long)	2HB Ticonderoga® brand pencil: Diameter: 0.2 inches, Shape = Hexagonal flat edges
Spice bottle (empty)	McCormick® brand Basil Leaves: Capacity = 0.62 Oz.
Teaspoon (metal)	Stainless steel teaspoon: Length = 5 15/16 inches, Weight = 0.05 lbs.
Writing paper pad	Ampad® brand gold fiber white ruled paper pad: Length = 11 3/4 inches, Width = 8 1/2 inches.

Note: No brands are endorsed in this measure and substitution is permitted keeping similar object characteristics (e.g., height, weight, texture, etc.)

Table 4.4. Characteristics of Participants with Neurological Conditions

Participant	*Diagnosis*	*Age*	*Sex*	*Education*	*Affected hand*	*Current Dominance*	*MAM ***	*Fatigue****	*Depression****	*Pain****
1	Stroke	60	F	Bachelors	R	R	90	45.8	55.7	55
2	Stroke	70	F	Bachelors	R	R	52.4	52.2	55.7	53.8
3	Stroke	65	M	PhD	L	L	66.4	52.2	54.5	52.5
4	PD	63	F	Bachelors	R	R	67.6	47.6	62.1	52.5
5	PD	61	M	Bachelors	R	R	63.3	53.7	53.3	41
6	PD	62	M	High School	R	R	59.6	62	57.9	41
7	PD	54	M	Not reported	L	L	53.9	50.8	56.8	56.1
8	Stroke	52	F	High School	L	R*	55.5	50.8	62.1	41
9	Stroke	46	M	Bachelors	L	L	49.3	43.9	64.1	41
10	Stroke	60	F	Masters	Both	Both	59.6	55.1	70.3	56.1
11	MS	50	M	Bachelors	L	R	68.9	53.7	53.3	41
12	MS	52	M	Bachelors	Both	R	66.4	47.6	65.1	52.5
13	MS	66	F	Bachelors	Both	R	52	66.3	60	61.8
14	PD	65	F	Masters	Both	R	54.7	60.6	61.1	60.9
15	PD	72	M	High School	Don't know	R	100	49.2	54.5	41
16	MS	51	M	Not reported	Both	R	49.3	67.8	68.2	48.5
17	MS	50	F	Bachelors	L	R	60.5	41.9	53.3	41
18	TBI	29	M	Less than High School	Both	R	47.8	55.1	70.3	59.1
19	MS	51	M	Bachelors	R	R	68.9	53.7	53.3	41
20	MS	51	F	Bachelors	L	R	60.5	41.9	53.3	41
21	Stroke	32	F	Bachelors	R	L*	65.3	43.9	53.3	52.5
22	PD	63	M	High School	R	B	59.6	62	57.9	41
23	Stroke	47	M	Bachelors	L	R	49.3	43.9	64.1	41
24	PD	62	M	Bachelors	L	R	63.3	53.7	53.3	41
25	TBI	30	M	Less than High School	Both	R	47.8	55.1	70.3	59.1
26	TBI	35	M	Masters	R	L*	63.3	50.8	64.1	55

Note: *Changed dominance due to hand problems. **Higher scores indicate more of the trait and thus better hand function. ***Higher scores on these Patient Reported Outcome Measure Information System (PROMIS) measure indicate more of the trait, thus higher scores indicate more fatigue, more depression and more pain. MAM: Manual Ability Measure-20; Fatigue: PROMIS Fatigue short form 7a version 1.0; Depression: PROMIS Depression short form 8b version 1.0; Pain: PROMIS Pain Interference short form 6b version 1.0

Table 4.5. New Items from Cognitive Interviews

Items added from cognitive interviews	
2 liter coke bottle lifting and pouring	Peeling an apple
Button with a loop, pin for a badge	Picking up different coins
Cards and toys	Picking up heavy items not in a bag
Carpenter work	Picking up marbles
Chopping jalapenos	Picking up things from floor
Close a paper clip	Pin in a pin cushion
Close a safety pin	Placing objects at different locations
Collar tabs	Pouring vanilla onto teaspoon
Construction work	Puppets
Cutting paper	Putting away groceries
Cutting tablets into half	Putting kids arms in coats
Dishes	Putting toys away
Drawing circles	Sample shirt buttoning
Dump something from a big pot to a colander	Scooping with a spoon
Filing	Screws and screwdriver use
Finger exerciser with rubber bands	Serving food
Flipping an egg	Slicing and dicing to the same size of vegetables
Flipping pages of a book	Small tablets
Flossing	Stacking cans
Folding papers	Stapling
Hammer	Threading the nut on a bolt
Holding big phones	Toothpaste on toothbrush
Holding book to read	Turning a pork chop frying on a pan
Holding kids hand	Tying a tie
Keyboard	Tying trash bags
Keys out of the pocket	Using a can opener
Make piles of 10 coins	Using a needle
Measuring with measuring spoon	Using a spatula
Moving big cans of tomato juice	Using a tape and mark a line
Opening a can of beans	Using pliers
Opening good wrapped in plastic	Using saw and cutting wood
Paper clips	Ziploc® bags
Passing dishes or cups	Zipping coats
Peel a kiwi	

Figure 4.1. Materials Used for Performance-based Section of the Measure

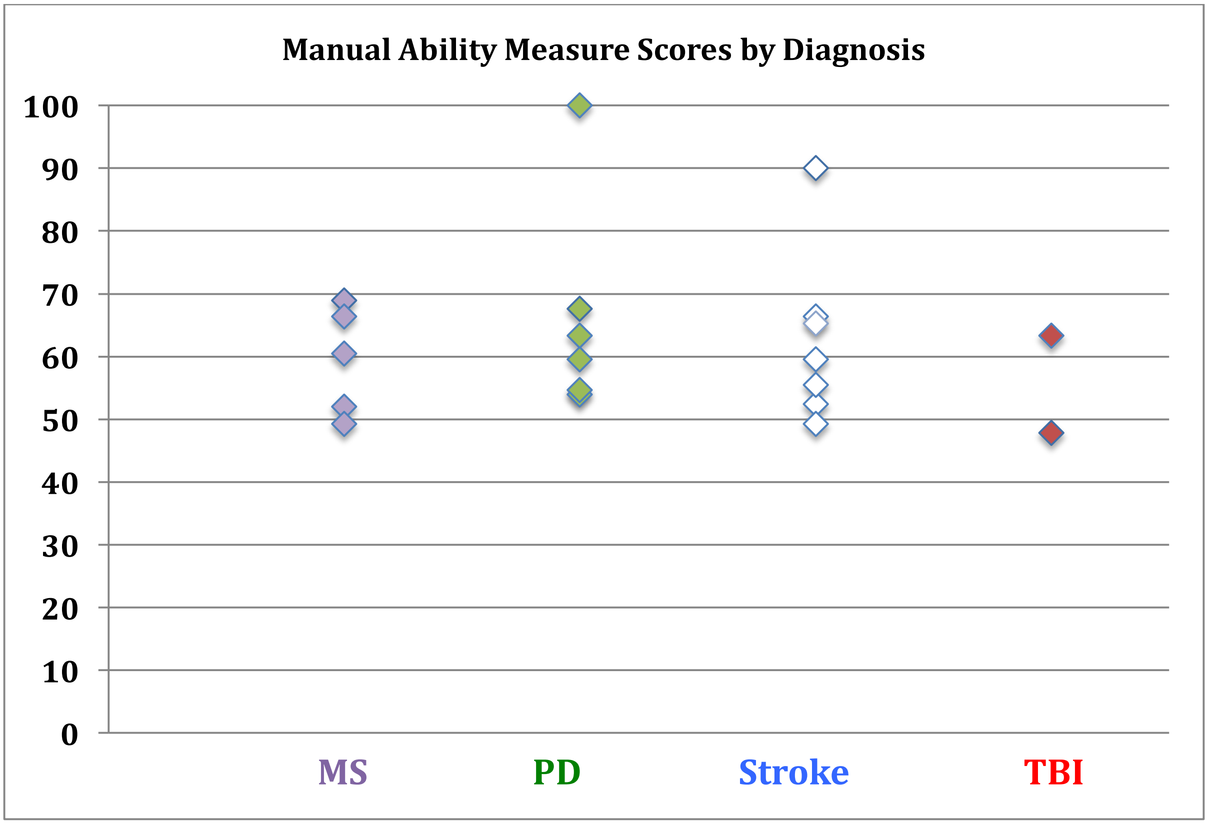


Figure 4.2. Manual Ability Measure Scores by Diagnosis

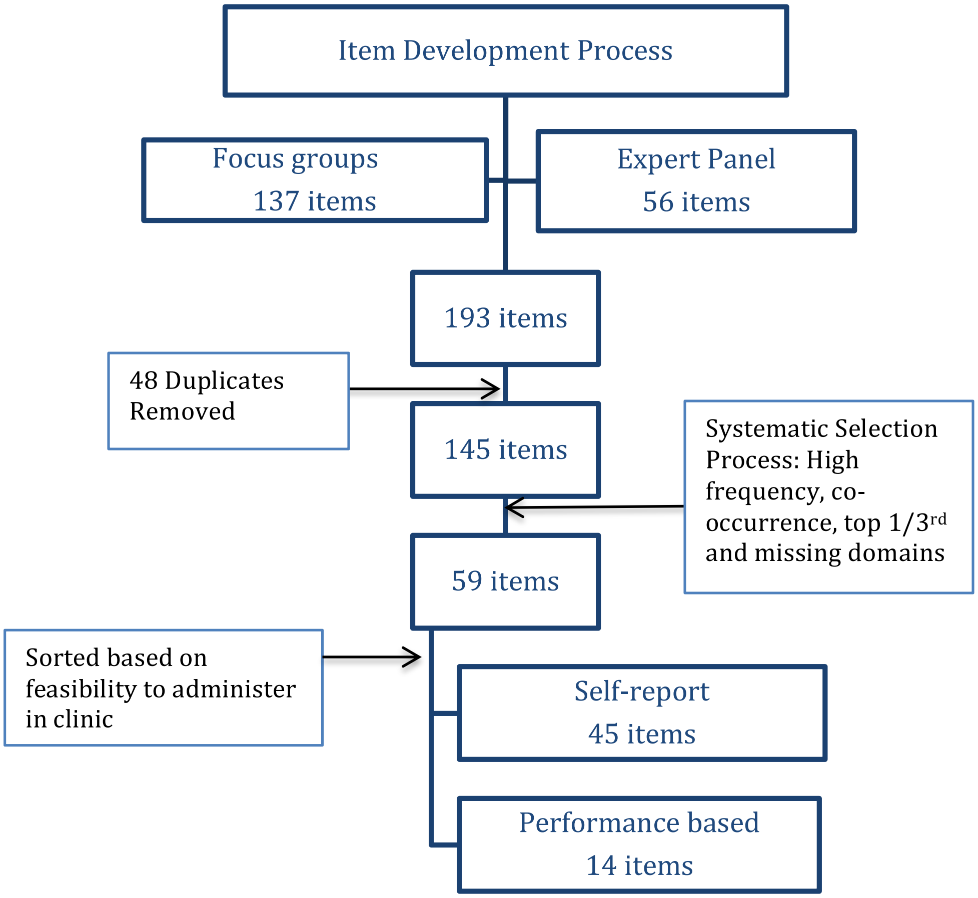


Figure 4.3. Item Development Process: Schematic Flow Diagram from Item Bank Development to Preliminary Selection of Items

Chapter 5: Discussion and Future Research

This dissertation represents the first phase of development of a new measure. The purpose of the Hand and Arm Function Measure (HAFM) is to assess the quality and speed of task performance and self reported difficulty with body structure and function, activity and participation due hand and arm function issues in daily activities for people with neurological conditions. These aspects are assessed through self-report and performance-based sections of the measure. This chapter summarizes the initial development, provides an overarching discussion, and presents a plan for future development of the HAFM.

Summary

Chapter 1 provided an introduction to hand function in general and hand function issues in particular in neurological conditions. A brief overview of the different types of hand function tests and their limitations were presented. The recognized need for a new measure of hand and arm function, addresses two key components, self-reported hand function and quality and speed of task performance. None of the currently available measures assess both of these components and have limited evidence for validity of scores for use in neurological conditions. In order to develop the new measure the framework of evidence-centered design (ECD) was adopted. The reasoning behind the selection of this framework was its utility in developing self-reported and performance-based sections, its emphasis on analysis of the construct and strong connection of conceptual framework with psychometric testing decisions. This chapter described the five layers of ECD, the purpose of the research project, and objectives of the dissertation research study.

Chapter 2 included a literature review of upper extremity function classification, neural correlates and motor control basis, a systematic review of measures of upper extremity function and limitations of the currently available measures. The classification by Jones and Lederman

(1993) was adopted for classifying the various functional aspects of hand and arm function. The model was selected for its focus on sensory motor function and was presented in detail in this chapter. The literature review on neural control provided a foundation for understanding the neurological control and motor control of hand and arm movements. The in-depth systematic review of hand and arm function tests presented in this chapter described 22 measures reported in the literature in the last ten years and discussed their limitations.

The development of the HAFM was informed by a panel of experts, focus groups, and cognitive interviews with people with neurological conditions. All provided useful insights into the definition of the construct, the structure of the measure, and the selection of items and responses. In Chapter 3, the construct of 'hand and arm function in daily activities' was explained with the help of a construct map and nomological network for neurological conditions. High level hand and arm function was defined as the ability of a person to participate in a wide range of daily activity tasks with the use of their hand and arm with optimal speed and quality, without experiencing interference by body structure and function problems and corresponds with higher scores on the HAFM. The International Classification of Functioning Disability and Health (ICF) framework that shaped this construct and the ECD framework that guided scale development are discussed in detail for their influences on the conceptualization of HAFM. The first phase of the research project involved engaging experts in defining the construct, item bank development, and item writing and modification. The item writing and modification were conducted after the focus groups with people with neurological conditions were conducted. Using a focus group format, the expert panel qualitative study was conducted from which eight key themes emerged. These themes informed the construct definition, item generation, shaped response format and scoring criteria, and identified the hand function issues faced by people with stroke, traumatic brain injury, multiple sclerosis and Parkinson disease and were described in detail in this Chapter 3.

For the next stage of the development of the HAFM, focus groups and cognitive interviews with people with neurological conditions were conducted and are presented in Chapter 4. The qualitative descriptive study results, their influence on the construct, item bank development, and item modification are discussed. The focus groups helped identify the aspects of hand function that were relevant to the participants with neurological conditions (i.e., sensation, strength, grasp, and fine hand use). The feedback from participants in each of the four diagnostic conditions helped to develop the item bank. The final item set was developed and preliminary items were administered to participants in the form of one-on-one cognitive interviews. Based on the cognitive interviews it was found that self-report and performance-based items were well tolerated by the participants. In addition, during the cognitive interviews, participants suggested modifications to the items, response options, and order of the items.

Discussion

The confidence in the results of this research emerge from triangulation. Triangulation is a strategy in qualitative research where the conclusions are drawn after data is gathered from different respondents, methodological approaches and frameworks (Golafshani, 2003). According to Golafshani (2003), there are four types of triangulation presented in literature, theoretical, methodological, data, and investigator. Theoretical triangulation relates to incorporating different theoretical frameworks. Methodological triangulation relates to having more than one methodological approach for collecting data. Data triangulation is concerned with data being gathered from more than one resource or participant group. Investigator triangulation occurs where two or more researchers take part in the data analysis. In this research study, theoretical triangulation was implemented with consideration of the ICF framework and Jones and Lederman's classification of hand function. Methodological triangulation included having more than one approach to explore the development of a new hand and arm function measure and included literature review, an expert panel, focus groups, and cognitive interviews. Data triangulation was addressed by collecting data from experts and participants with neurological conditions. Investigator triangulation was addressed by having with two researchers conduct the data analysis. These triangulation strategies have strengthened the qualitative approach utilized for developing the HAFM to inform face and content-related evidence for validity of the scores of the measure. The quantitative perspective of triangulation for item development is evident in the methodology of this research with a lack of deep discovery of the phenomenon of hand function problems. In recent years, similar methods have been utilized by other researchers in developing item banks (Morgan et al., 2014).

Limitations. This research study has limitations in the selection of one geographic location for data collection. It is also limited by small sample size, particularly for TBI. Some scale development studies employ large pool of participants (and include international participants (Annear et al., 2015). Although the current study has a small sample size, the participants were people with neurological conditions who are more difficult to recruit in clinical trials. This is also a limitation for applying the state-of-the-art Item Response Theory (IRT) methods since they require a large sample size for accurate statistical analysis. The fit statistics and differential item functioning cannot be conducted to establish construct-related evidence for validity of scores without the use of IRT based approaches. Although the Classical Test Theory (CTT) methods are not as robust as the IRT methods, they are the statistical method of choice for the current planning of the pilot study.

Another limitation in the sample is the focus on four diagnostic conditions of stroke, traumatic brain injury (TBI), multiple sclerosis (MS), and Parkinson disease (PD). The measure, in its current form, is not designed for use with people with other types of neurological conditions. The studies of evidence for the reliability and validity of scores would need to be conducted in the future for use with people with these and other types of neurological conditions. The plethora of neurological conditions that could result in hand and arm function problems make it impossible to include all types of diagnoses in the study population. The two non-progressive conditions of stroke and TBI and the two progressive conditions of MS and PD were assumed to be representative of the most commonly seen neurological conditions in the clinic that affect hand and arm function. Other conditions such as spinal cord injury (SCI) that can affect hand function were excluded from this study since the type of hand function problems in SCI are very unique to the condition (e.g., proximal more than distal issues in central cord syndrome), change considerably with the level of SCI (e.g., person with C6 tetraplegia with lost triceps function that is preserved in C7 tetraplegia), and involve special compensatory techniques (e.g., tenodesis in C6 tetraplegia). Moreover, specific measures such as the Capabilities of Upper Extremity (CUE) (Marino, Kern, Leiby, Schmidt-Read, & Mulcahey, 2015) and Grasp and Release Test (GRT) (Mulcahey, Smith, & Betz, 2004) have been developed in self-report (CUE) and performance-based (CUE and GRT) formats to measure hand and arm function in people with SCI. There are potentially many other neurological conditions, for which the measure would need to be validated.

Future Research

The next step in the development of the HAFM is the pilot administration of the measure. The results of pilot administration can influence the construct, design of the measure, item design, and item bank.

HAFM Pilot Study. A brief blueprint for pilot administration is subsequently presented.

Purpose. The objective of the HAFM pilot study will be to investigate internal consistency, convergent, and discriminant evidence for validity of scores and to revise items as needed. The pilot study will involve administering the test items to people with neurological disorders such as stroke, TBI, MS, PD and people who do not have neurological disorders. The quantitative study will involve administering the HAFM and five other standardized measures of hand and arm function, and conducting preliminary psychometric analysis using classical test theory methods. The qualitative study will involve conducting 'think alouds' at the end of the complete administration of measure to get participant feedback on structure, flow, order, time and instructions related to entire measure and particular items revised in this iteration of the measure.

Participants. Forty adult participants, with eight participants in each group (stroke, TBI, MS, PD, and no neurological condition) will be recruited. Ideally, participants will represent a diverse range of age, ethnicity, and stage of the neurological condition (as appropriate) and have an equal representation of males and females. It is desired that these individuals have varying severities of hand and arm function from mild to severe and healthy adults. After human subjects approval, participants will be screened using a telephone interview. A hand function questionnaire such as the Stroke Impact Scale – Hand Domain (SIS-HD) (Duncan et al., 2002) will be used for screening. Participants will be enrolled in the study after inclusion criteria have been met through the screening and written informed consent is obtained.

Procedures. Each data session will consist of obtaining demographic information and administering a battery of assessments. Assessment will include the HAFM, the Montreal Cognitive Assessment Basic (Freitas, Simões, Marôco, Alves, & Santana, 2012; Nasreddine et al., 2005), the National Institute of Health (NIH) Patient Reported Outcomes Information System (PROMIS) – measures for depression, fatigue and pain interference short forms (Gershon, Rothrock, Hanrahan, Bass, & Cella, 2010), the Upper Extremity Neurological Quality of Life Short form (Cella et al., 2012), the Manual Ability Measure - 20 (Rallon & Chen, 2008), the Quick DASH (Veehof, Sleegers, van Veldhoven, Schuurman, & van Meeteren, 2002), the Nine Hole Peg Test (Oxford Grice et al., 2003), hand strength assessment for grip and pinch with the dynamometer (Peters et al., 2011), the Jebsen Hand Function Test (Stamm, Cieza, Machold, Smolen, & Stucki, 2004) and the Timed Up And Go test (Mesquita et al., 2013). The self-report section of HAFM will be administered before the performance-based section.

Preliminary Data Analysis Plan. Classical item analysis statistics will be utilized for data analysis for this pilot study (Allen & Yen, 2002). The statistics package Winsteps® (www.winsteps.com) will be used for analysis. This package was selected for its statistical tools that provide measurement related analysis. A multimethod multitrait matrix will be constructed as shown in Table 5.1. It is expected that the fine movements items on both sections of the HAFM will correlate highly and so will the gross movements items. To measure internal consistency, Cronbach's alpha will be computed for the test and acceptable scores are considered to be 0.70 or above and strong scores are considered to be 0.90 and above (McDowell, 2006). Convergent evidence for validity of scores can be established separately for each section of the measure since there is no other standard measure that has both sections in one measure. The relationship between the performance-based section of the measure, the Nine Hole Peg Test, the Jamar® dynamometer, and the Jebsen Hand Function Test; and the self-report section of the measure with the UE NeuroQoL Shortform, the Quick DASH and the MAM-20 will be explored. A strong correlation is expected between performance-based section and JHFT, and self-report section, the UE NeuroQoL Shortform, and the MAM-20. A moderate correlation is expected between performance-based section of the measure with NHPT and dynamometer, and the self-report section with DASH. The discriminant evidence for validity of scores will be obtained by examining the relationship between HAFM and TUG (Timed Up and Go Test) will be examined with the hypothesis that there is expected to be a low correlation. A low correlation is expected versus no correlation because participants who struggle with mobility may be the more severely affected individuals; therefore, their hand and arm function may also be impacted. However, the correlation between TUG and the scores from the new measure should be lower than the correlations between scores from other hand function measures and scores from the new measure.

The construct-related evidence for validity of scores cannot be established by one study and has to be a combined result of multiple studies with varied samples (American Educational Research Association, 1999). In the recent literature related to hand and arm function assessment measures, as noted in the systematic review in Chapter 2, many studies in development of a measure are incomplete and do not address all psychometric properties, leaving the clinician with the choice of informal testing versus a poorly studied measure. Thus, the scope of this project is outlined to raise the standards of evidence to support validity and reliability of scores of the clinical measure of hand and arm function.

Scope of research project. The overall goal of this project is to develop a reliable and valid measure of hand and arm function in people with neurological conditions that is psychometrically robust. Nine phases of the overall scope of this project were developed without considering the limitations of resources, participants and personnel and the research plan is outlined in Table 5.2. The item bank development with stakeholder feedback using focus groups and cognitive interviews is completed with evidence for content established in the current study. The results from the current study and the proposed pilot study will be used to help secure funding for larger studies. The item tryout study will recruit a representative sample (N = 100) for standardized administration of the measure and run classical test theory (CTT) statistics to compute item level statistics. Item level statistics will help determine which items have a range of popularity (self-report section) or ability (performance-based section) and if all items contribute to their total score. The revised measure will then be ready for large-scale administration where a stratified sample (N = 400) of participants across different settings and locations is expected for psychometric analysis. The state-of-art psychometric methods of item response theory (IRT) can be applied to this large sample to determine internal consistency, convergent and discriminant evidence for validity of scores and establishing cut off scores. Cut off scores will be determined using the contrasting groups method (Livingston & Zieky, 1982). In this method, the therapists will be asked to determine severity of hand function deficits into categories of mild, moderate and severe. The cut-score range that corresponds to these categories will be computed. The healthy adults will provide the unimpaired category scores. Random samples can be pulled from this large sample for test retest and inter-rater evidence for reliability of scores, and convergent and discriminant evidence for validity of scores. Evidence to support the intended use of scores can then be investigated in a randomized controlled trial where the pre and post scores are compared for clinically meaningful change and minimal clinically important difference. A longitudinal follow-up of a subset of participants to gather evidence for unintended consequences of the measure is planned after the standardized measure is studied for clinical use.

Conclusion

To conclude, in the words of Robert DeVellis - a renowned measurement expert famous for his work with the NIH initiative on PROMIS and more than 30 years of experience in scale development, - *"Even if a poor measure is the only one available, the costs of using it may be greater than any benefits attained."* (DeVellis, 2012) (p. 10)

Clients and therapists should not have to spend valuable clinical time on measures that are poorly studied and likely of limited benefit to the client. Stakeholder engagement in the initial stages is key to developing the construct and the items that are valid for measuring hand and arm function in people with neurological conditions. Patient/client reported hand and arm function is valuable in measuring outcomes of rehabilitation. The HAFM is grounded in these concepts and is unique in its combination of self-reported and performance-based aspects included within a single measure.

This research has focused on the preliminary development of the HAFM with attention to rigor in designing a measure that reflects the needs of patients/clients and clinicians. This is a first step in the development of the HAFM. Much more additional work is necessary, from piloting to large-scale administration, for ongoing development to gather evidence for validation of scores of the measure.

Table 5.1. Multitrait Multimethod Matrix with Expected Correlations

		HAFM Performance-based section		*HAFM Self-report section*	
		Fine movements	**Gross Movements**	**Fine movements**	**Gross Movements**
HAFM Performance-based section	**Fine movements**	1.00	Moderately High	High	Moderately High
	Gross movements	Moderately High	1.00	Moderately High	High
HAFM Self-report section	**Fine movements**	High	Moderately High	1.00	Moderately High
	Gross movements	Moderately High	High	Moderately High	1.00

Table 5.2. Planning for Research to Support Interpretation and Use of Assessment Scores

Purpose	*Sample*	*Population*	*Analyses*	*Rationale*	*Status*
Item bank development (Phase I)	N = 16 participants, focus groups N = 8 participants, cognitive interviews Purposeful sampling	Participants with neurological conditions of stroke, TBI, MS, and PD. Outpatient setting. Equal gender representation, diversity among age and ethnic background	Qualitative analysis • Ask participants to list difficulties with hand and arm function in daily activities. • Ask participants to respond to items and write comments about clarity of items. • Ask participants whether the tasks and wording are universal and culturally appropriate. • Ask about parts that are unclear including scoring	• Gather lived experiences with hand and arm problems of people with neurological condition to build content-related evidence for validity of scores • Pre-piloting of items for appropriateness with the 'think aloud' method.	Done
Content-related evidence for validity for evidence (Phase I)	N = 4 Expert Panel members Purposeful sampling	Expert panel to include psychometrician and therapists who are subject matter experts in neurological rehabilitation in clinic, teaching and research.	Qualitative analysis • Ask about the clinical needs, construct definition • One round of item bank development. • Second round of item checking where the experts are shown the items currently in the measure and their opinion about items related to the construct is reported	• Expert panel can help us identify items that are best suited for the measure • The second round is when the experts are shown the items and their comments are used to revise the items	Done
Pilot administration (Phase II)	N = 40. Convenienc e sample with 8 participants in each of the 4 subgroups of neurological conditions and healthy adults	Participants with neurological conditions of stroke, TBI, MS, PD, and healthy adults. Equal gender representation, diversity among age, hand function and ethnic background. Clinic based administration in inpatient or outpatient settings.	Mixed method analyses • Administer the HAFM and other standardized measures. • 'Think alouds' done at the end to comment on the entire test, the order of tasks and the time for entire test completion. CTT methods used for: • Internal consistency • Convergent evidence for validity of scores • Discriminant evidence for validity of scores Revision of standardization protocols and items based on results.	• New and revised items from cognitive interviews piloted. • Order effects for presentation of self-report and performance-based sections to be explored. • Norms for the timed tasks using data from healthy adults and people with neurological conditions is gathered to categorize the timed tasks and inform scoring. • Secure funding for large scale studies	Planned

Table 5.2. Planning for Research to Support Interpretation and Use of Assessment Scores (Continued)

Purpose	*Sample*	*Population*	*Analyses*	*Rationale*	*Status*
Item tryouts (Phase III)	N = 100. Convenience sample with 20 participants in each subgroup of neurological conditions and healthy adults.	Participants with neurological conditions of stroke, TBI, MS, PD, and healthy adults. Participants from inpatient and outpatient settings. One rural, one sub-urban and one urban setting is preferred.	Quantitative analyses using CTT methods. • Administer the HAFM • Compute item statistics: o Item means o Item to total score correlations for the two formats o Item factor analyses o Internal consistency o Convergent/ discriminant evidence for validity of scores	• Check to see if items have a range of popularity (self report) or ability (performance) with the help of item means • All items contribute to their total score.	Pending
Large scale testing: Standardization and norming (Phase IV)	Sample of 400 participants across different locations. 80 participants in each of the sub-populations.	Stratified sample of participants with: Layer 1 - neurological problems with hand function impairment at different levels, healthy adults. Layer 2 - inpatient and outpatient settings. Layer 3 - One rural, one suburban and one urban setting desired. Layer 4 - Equal males and females are desired. Layer 5 - Cultural mix is desired.	Quantitative analyses using IRT methods. • Administer the HAFM and other measures to representative sample. • Compute subscale score statistics of: Score means and standard deviations, score frequency distributions, score for different neurological conditions, Score for males and females, factor analyses for dimensions • Alpha coefficient • The therapists administering the HAFM will be asked to determine mild, moderate and severe hand function impairment based on their judgment and the contrasting groups method will be used to determine cut scores.	• Use IRT to create interval scale • Determination of clinically meaningful change • Determination of cut scores for mild, moderate and severe impairments based on contrasting groups method. • Dimensionality – check for the two dimensions of gross movements and fine movements for performance-based and symptoms and daily activities for self-report section. • Normative data for timed tasks.	Pending

Table 5.2. Planning for Research to Support Interpretation and Use of Assessment Scores (Continued)

Purpose	*Sample*	*Population*	*Analyses*	*Rationale*	*Status*
Test Retest Evidence for reliability of scores (Phase V)	N = 100 patients with 20 in each sub-population. A random sample from the standardization sample may be pulled out for evidence for reliability of scores.	Participants with neurological conditions of stroke, TBI, MS, PD, and healthy adults. Different levels of hand function impairments in the mild, moderate and severe categories	• Administer HAFM and other measures twice within one week interval between testing • Compute correlations between scores from first and second testing events.	• Alpha coefficients provide information if participants respond consistently on items within a subscale; • Test-retest evidence for reliability of scores informs if participant scores are consistent over time	Pending
Inter-rater Evidence for reliability of scores (Phase VI)	N = 100 patients with 20 in each sub-population. Applies to performance-based section of the measure.	Participants with neurological conditions of stroke, TBI, MS, PD, and healthy adults. Different levels of hand function impairments in the mild, moderate and severe categories	• Use video recordings from standardization sample, shown to two raters and their responses recorded. • Two raters administer the HAFM and their responses recorded.	• The performance format of the test should be consistent when administered by two raters. • High ICC between the two raters are desired for the measure when scoring the responses.	Pending
Construct-related evidence for validity (Phase VII)	N = 100 patients with 20 in each sub-population.	Participants with neurological conditions of stroke, TBI, MS, PD, and healthy adults By comparing healthy adults and persons with impaired hand function, differences between them can be explored.	Quantitative analyses • Administer HAFM and other measures such as MAM-20 and JHFT. • No gold standard is available, and since the HAFM self-report section is not a precise reflection of the construct presented in the MAM-20, only moderate correlations are expected. • JHFT only measures time to completion and in comparison with time component of HAFM, moderate correlations are expected here as well.	• MAM-20 correlations with the self-report section of HAFM will indicate if we are tapping into a similar construct. • JHFT is a performance-based measure that is generic and will tell us if we are measuring our performance in comparison with this measure. • Known group evidence for validity of scores can be explored by comparing healthy adults with participants with neurological conditions	Pending

Table 5.2. Planning for Research to Support Interpretation and Use of Assessment Scores (Continued)

Purpose	*Sample*	*Population*	*Analyses*	*Rationale*	*Status*
Evidence to support intended use of scores (Phase VIII)	N = 120 Two groups of individuals with neurological conditions randomized into control (N = 60) and treatment (N = 60) conditions.	Participants with neurological conditions of stroke, TBI, MS, PD	Quasi experimental design of quantitative study • HAFM administered to patients before and after an intervention in the treatment group and the control group. • Therapists set the mild, moderate and severe categories based on performance and these classifications are matched with the cut scores obtained. • Also the pre-post scores are compared for clinically meaningful change.	• To investigate if the test is responsive to detect change with intervention, specificity to not change with no intervention and ability to classify individuals into categories of impairment level as mild, moderate and severe.	Pending
Evidence for consequences (Phase IX)	N = 100 patients with 20 in each sub-population. A subset of people who participated in the treatment condition will be followed over time.	Participants with neurological conditions of stroke, TBI, MS, PD	Qualitative study • Follow the change in function over time in progressive and non-progressive conditions for about 2 years. • Intended consequences: Did the therapists use the results to plan interventions and do they relate to the scores? Did the therapists continue to use the measure to track progress? Assess how insurance reimbursement was tied to the reporting of scores on HAFM. • Unintended consequences: Did the measure create negative perceptions of self worth? Did the therapist's bias in some form create an inaccurate scoring of the measure? Construct irrelevant variance: Did another factor such as depression affect the performance of the individuals on the HAFM. • Ask the therapists about how they interpreted the test scores and what did they use it for. Assess how insurance reimbursement was tied to the reporting of scores on this measure.	• Longitudinal perspective is important to find out how patients and therapists used the measure and how the measure affected their perspective towards hand function or the test effect.	Pending

Note: CTT: Classical Test Theory, HAFM: Hand and Arm Function Measure, ICC: Intra Class Correlation, IRT: Item Response Theory, JHFT: Jebsen Hand Function Test, MAM-20: Manual Ability Measure – 20, MS: Multiple Sclerosis, PD: Parkinson disease, TBI: Traumatic Brain Injury.

Appendix A. Definitions for Terms Used in this Study Related to Measurement Criteria

Term	*Definition.*
Alpha	Also known as the Cronbach's alpha. It is a generalized formula used to express the internal consistency evidence for reliability of scores of a measure and varies from 0 to 1.
Ceiling effects	The effect in a measure where a high proportion of participants in a study have maximum scores.
Claims and evidence worksheets	The claims made by an assessment tool and the evidence that supports that claim are presented in the claims and evidence worksheets. (Mislevy et al., 2003)
Concurrent evidence for validity of scores	Evidence for validity of scores indicated by comparing scores on a measurement with those obtained by applying alternative, equivalent measurement at the same time.
Content-related evidence for validity of scores	The degree to which a measurement covers all aspects of the topic it purports to measure.
Convergent evidence for validity of scores	The degree to which two or more instruments that purport to be measuring the same topic agree with each other.
Correlation	The degree to which to which two or more sets of observations fit a linear relationship. It lies between -1 to +1.
Criterion-related evidence for validity of scores	The degree to which the results obtained using a measurement scale agree with a criterion standard or indicator of true situation.
Cross cultural evidence for validity of scores	The degree to which two or more groups of culturally diverse groups agree with each other.
Design patterns	Design patterns are tools to organize assessment arguments and are meant to guide families of assessment tasks around aspects of proficiency (Mislevy et al., 2003).
Discriminant evidence for validity of scores	The extent to which scores on a measurement distinguish between individuals or populations that would be expected to differ (e.g., people with or without stroke)
Floor effects	The lower limit to scores that cannot be reliably reported and people with low level of the trait cannot be distinguished, is known as the floor effect.
Inter-rater evidence for reliability of scores	The extent to which the results obtained by different raters or interviewers using the same measurement method will agree.
Internal consistency	It is the degree to which all the items in a test, measure the same theme.
Intra-class correlation coefficient	The degree of agreement that records the average similarity of rater's actual scores on the ratings being compared.
Intra-rater evidence for reliability of scores	The extent to which the results obtained by the same rater at two different time points using the same measurement method will agree.
Minimal clinically important difference	The smallest change on a health outcome measure that would be considered important by patient or clinician.
Minimal detectable change	The smallest change on a health outcome measure that corresponds to a noticeable change in ability.
Predictive evidence for validity of scores	The ability of the outcome of an instrument can predict a future outcome.
Evidence for reliability of scores	The proportion of variance in a measurement that is not related to error.
Evidence for responsiveness	The ability to detect change, critical for outcome measures.
Standard error of measurement	The standard deviation of errors, the amount of error due to chance.
Test-retest evidence for reliability of scores	The repeatability of measurement evaluated in terms of agreement between the measure administered at two different time points.
Toulmin diagrams	Toulmin diagrams are examples of tools for organizing assessment arguments at a narrative level (Mislevy et al., 2003).
Evidence for validity of scores	The extent to which a measure is measuring what it was intended to measure.

Note: Definitions adapted from the book 'Health Measurements'(McDowell, 2006) except when indicated otherwise

Appendix B. Studies Identified in the Systematic Review for the Self-report Measures

Measure	*Studies Identified in the Systematic Review*
Disabilities of Arm Shoulder and Hand	3 studies (Cano, Barrett, Zajicek, & Hobart, 2011; Hijmans, Hale, Satherley, McMillan, & King, 2011; Kraft et al., 2014)
ABILHAND	2 studies (Alt Murphy, Willén, & Sunnerhagen, 2012; Kraft et al., 2014)
Motor Activity Log	17 studies (Bowman et al., 2006; Brogårdh & Lexell, 2010; Brunner, Skouen, & Strand, 2012; Celik et al., 2010; Dahl et al., 2008; Hammer & Lindmark, 2010; Lamers, Kerkhofs, et al., 2013; Lamers, Timmermans, et al., 2013; Lin, Wu, Wei, Lee, & Liu, 2007; Keh-chung Lin, Chang, Wu, & Chen, 2009; Lin, Huang, Hsieh, & Wu, 2009; Pang, Harris, & Eng, 2006; Taub et al., 2013; Timmermans et al., 2014; van Delden, Peper, Beek, & Kwakkel, 2013; Steven L. Wolf et al., 2006)
Stroke Impact Scale – Hand Domain	11 studies (Boyne, Dunning, Levine, Hermann, & Page, 2010; Combs, Kelly, Barton, Ivaska, & Nowak, 2010; Dahl et al., 2008; Gurcay, Bal, & Cakci, 2009; Huang, Wu, Hsieh, & Lin, 2010; Keh-chung Lin, Chang, et al., 2009; Keh-chung Lin, Wu, et al., 2009; Page et al., 2011; L. Shaw et al., 2010; van Delden, Peper, Beek, et al., 2013; Winstein et al., 2013)
Manual Ability Measure	1 study (Kraft et al., 2014)
Neurological Quality of Life – Upper Extremity	1 study (Kraft et al., 2014)
Patient Reported Outcome Measure Information System – Upper Extremity	2 studies (Kraft et al., 2014; Lamers, Kerkhofs, et al., 2013)

Appendix C. Studies Identified in the Systematic Review for the Performance-based Measures

Measure	*Studies Identified in the Systematic Review*
Jebsen Hand Function Test	4 studies (Adamovich et al., 2009; Celik et al., 2010; Kraft et al., 2014; Wu, Seo, & Cohen, 2006)
Rivermead Movement Assessment – Upper Extremity	1 study (J. H. Morris et al., 2008)
Motricity Index – Upper Extremity Section	3 studies (Lamers, Kerkhofs, et al., 2013; Lamers, Timmermans, et al., 2013; L. Shaw et al., 2010)
Action Research Arm Test	34 studies (Alt Murphy et al., 2012; Au-Yeung & Hui-Chan, 2014; S. R. Barreca et al., 2005; Boyne et al., 2010; Brunner et al., 2012; Burns, Burridge, & Pickering, 2007; Cano et al., 2011; Celik et al., 2010; H. Chen, Lin, Wu, & Chen, 2012; Chuang, Wu, & Lin, 2012; Hammer & Lindmark, 2010; Hughes et al., 2009; Ietswaart et al., 2011; Kraft et al., 2014; Lamers, Kerkhofs, et al., 2013; Lamers, Timmermans, et al., 2013; J.-H. Lin, Hsu, et al., 2009; McDonnell, Hillier, Ridding, & Miles, 2006; J. H. Morris et al., 2008; J. H. Morris & Van Wijck, 2012; Page, Hade, & Persch, 2015; Page, Levine, & Hade, 2012; Pomeroy et al., 2014; Rabadi & Rabadi, 2006; Rosewilliam, Malhotra, Roffe, Jones, & Pandyan, 2012; Seo, Fischer, Bogey, Rymer, & Kamper, 2011; L. Shaw et al., 2010; Timmermans et al., 2014; Turk et al., 2008; Urbin, Waddell, & Lang, 2015; van Delden, Peper, Nienhuys, et al., 2013; van Delden, Peper, Beek, et al., 2013; Wu et al., 2006; Yozbatiran et al., 2008)
Fugl-Meyer Assessment – Upper Extremity	34 studies (Alt Murphy et al., 2012; Arya & Pandian, 2013; Bartolo et al., 2014; Bowman et al., 2006; Celik et al., 2010; Chae et al., 2009; Chan, Tong, & Chung, 2009; Daly et al., 2005; Hammer & Lindmark, 2010; Hesse et al., 2005; Hijmans et al., 2011; Hughes et al., 2009; Knutson et al., 2012; Krebs et al., 2008; Lamers, Kerkhofs, et al., 2013; Lamers, Timmermans, et al., 2013; Levy et al., 2008; J.-H. Lin, Hsu, et al., 2009; Keh-chung Lin, Chang, et al., 2009; Keh-Chung Lin, Huang, et al., 2009; Keh-chung Lin, Wu, et al., 2009; Masiero, Celia, Rosati, & Armani, 2007; McDonnell et al., 2006; Ohn et al., 2013; Page et al., 2011, 2015, 2012; Pang et al., 2006; Rabadi & Rabadi, 2006; Reinkensmeyer et al., 2012; Rosati, Gallina, & Masiero, 2007; Sullivan, Hurley, & Hedman, 2012; Taub et al., 2013; Turk et al., 2008)
Function Test of Hemiplegic Upper Extremity	1 study (Chan et al., 2009)
Box and Blocks Test	6 studies (Higgins et al., 2006; Knutson et al., 2012; Kraft et al., 2014; Milot et al., 2014; Reinkensmeyer et al., 2012; Seo et al., 2011)
Nine Hole Peg Test	7 studies (Brunner et al., 2012; Higgins et al., 2006; Kraft et al., 2014; Lamers, Kerkhofs, et al., 2013; J. H. Morris et al., 2008; J. H. Morris & Van Wijck, 2012; Polman & Rudick, 2010; L. Shaw et al., 2010)
Function Test of Hemiplegic Upper Extremity	1 study (Chan et al., 2009)
Box and Blocks Test	6 studies (Higgins et al., 2006; Knutson et al., 2012; Kraft et al., 2014; Milot et al., 2014; Reinkensmeyer et al., 2012; Seo et al., 2011)

Appendix C. Studies Identified in the Systematic Review for the Performance-based Measures (Continued)

Hand Function Measure	*Studies Identified in the Systematic Review*
Function Test of Hemiplegic Upper Extremity	1 study (Chan et al., 2009)
Box and Blocks Test	6 studies (Higgins et al., 2006; Knutson et al., 2012; Kraft et al., 2014; Milot et al., 2014; Reinkensmeyer et al., 2012; Seo et al., 2011)
Nine Hole Peg Test	7 studies (Brunner et al., 2012; Higgins et al., 2006; Kraft et al., 2014; Lamers, Kerkhofs, et al., 2013; J. H. Morris et al., 2008; J. H. Morris & Van Wijck, 2012; Polman & Rudick, 2010; L. Shaw et al., 2010)
Motor Assessment Scale	4 studies (Brauer, Hayward, Carson, Cresswell, & Barker, 2013; Brogårdh & Lexell, 2010; Hammer & Lindmark, 2010; Sabari et al., 2005)
Sollerman Hand Function Test	1 study (Brogårdh & Lexell, 2010)
Arm Motor Ability Test	6 studies (Chae et al., 2009; Daly et al., 2005; Daly & Ruff, 2007; Knutson et al., 2012; Levy et al., 2008; Sullivan et al., 2012)
Wolf Motor Function Test	11 studies (Adamovich et al., 2009; Borstad et al., 2013; Bowman et al., 2006; Combs et al., 2010; Dahl et al., 2008; Hijmans et al., 2011; J.-H. Lin, Hsu, et al., 2009; Pang et al., 2006; Sawaki et al., 2008; Winstein et al., 2013; Steven L. Wolf et al., 2006)
Test Evaluant la Performance des Membres supérieurs des Personnes âgées	2 studies (Higgins et al., 2006; Kraft et al., 2014)
Stroke Rehab Evaluation Assessment of Movement	2 studies (Higgins et al., 2006; J.-H. Lin, Hsu, et al., 2009)
Chedokee Arm and Hand Inventory	6 studies (Abdullah, Tarry, Lambert, Barreca, & Allen, 2011; S. R. Barreca et al., 2005; Gustafsson, Turpin, & Dorman, 2010; Harris, Eng, Miller, & Dawson, 2009; Schuster-Amft et al., 2014; Wu et al., 2006)

Appendix D. Item Bank with Descriptions

Item	*Description*
Carrying: Carry groceries	Able to carry groceries and bring it home or take it to the car.
Carrying: Coffee cup	Able to carry a coffee cup and take it with you.
Carrying: Shoulder bag	Able to carry a bag on the shoulder
Carrying: Wine glass	Able to carry a wine glass in the affected hand.
Communication: Clapping	Able to clap hands
Communication: Shaking hand	Able to shake hands with another person
Communication: Writing	Able to use a writing device like pen or pencil and sign name, make lists in a legible handwriting within a reasonable amount of time.
IADL: Childcare tasks	Able to hold and carry a child, play with the child.
IADL: Drive thru services	Able to use the drive thru services (e.g. fast food, parking tickets).
IADL: Grocery: General	Able to do general grocery related tasks like going to grocery store.
IADL: Grocery: Opening thin plastic bags at store	Able to open thin plastic bags at the store to put grocery items in it like vegetables.
IADL: Grocery: Sign at cashier's counter	Able to sign your name on the receipt or on the electronic pad at the cashier's counter.
IADL: Hanging clothes on line	Able to hang clothes on a clothes line
IADL: Housekeeping: General	Able to do general housekeeping tasks
IADL: Housekeeping: Mopping	Able to hold a mop with or without a handle and mop the floor
IADL: Housekeeping: Pet care	Able to take care of the pet, stroking the pet, caring for the dog or cat.
IADL: Housekeeping: Vacuum	Able to vacuum the carpet.
IADL: Kitchen tasks: Cooking	Able to cook a meal.
IADL: Kitchen tasks: Cracking an egg	Able to break an egg without spilling or dropping the shell, able to get the shell out if dropped.
IADL: Kitchen tasks: Creamer cups	Able to open creamer cups that need small pinch and strength to pull
IADL: Kitchen tasks: Cutting	Able to cut up food in the kitchen like vegetables, meat.
IADL: Kitchen tasks: Getting plastic off the cheese slices	Able to get the plastic off the cheese slices
IADL: Kitchen tasks: Kneading	Able to knead the dough.
IADL: Kitchen tasks: Knife	Able to use a knife, cutting knife or pocket knife.
IADL: Kitchen tasks: Microwave	Able to use the microwave.
IADL: Kitchen tasks: Peeling	Able to hold the vegetable in one hand use a peeler with the other.
IADL: Kitchen tasks: Pick up pot on stove	Able to pick up a pot with hot contents from the stove.
IADL: Kitchen tasks: Plates	Able to hold and carry plates to and from the kitchen
IADL: Kitchen tasks: Pour water/soups/milk	Able to pour water, soup, milk
IADL: Kitchen tasks: Stirring	Able to stir hot items, particularly holding the container and stirring
IADL: Laundry, folding clothes	Able to fold clothes, using the washer/dryer
IADL: Laundry: Towel - folding	Able to fold a towel
IADL: Pushing - wheelchair, cart	Able to push the grocery cart, wheelchair for someone else with your hands
IADL: Washing the car	Able to wash own car
IADL: Work: General	Issues with work
IADL: Work: Return to work	Issues with return to work
IADL: Work: Type of work	Different types of paid and unpaid work
IADL: Work: Vocational Rehab	Able to get back to work using vocational rehabilitation services
Movement: Holding steady	Able to hold objects steady
Movements: Lifting	Able to lift items
Movements: Opening/Closing hand	Able to open and close the hand
Movements: Pick up	Able to pick up items

Appendix D. Item Bank with Descriptions (Continued)

Item	***Description***
Movements: Placing on table	Able to place hand on table
Movements: Tap on table	Able to tap on the table
Object handling: Bags, purses, backpack	Manipulating bags, purses, backpacks
Object handling: Ball	Able to catch or throw ball
Object handling: Bottle	Able to pick up, close and open a bottle and handle various caps
Object handling: Cans	Able to open, close, hold and pick up cans
Object handling: Childproof bottle for medications	Able to open and close childproof bottle for medications
Object handling: Coins	Able to gather coins, count them in hand
Object handling: Credit cards handling	Able to handle credit cards during shopping
Object handling: Cup	Able to pick up and pour into a cup
Object handling: Doors	Able to open and close doors.
Object handling: Filing papers	Able to file papers
Object handling: Hammer/Nail	Able to hammer a nail
Object handling: Jar opening	Able to open and close a jar
Object handling: Keys	Able to use keys to open doors, locks
Object handling: Kitchen tasks: Flipping	Able to flip like an omelet when cooking to get the other side cooked.
Object handling: Lids	Able to open and close lids, particularly the tight ones
Object handling: Light switches	Able to use light switches in the home and outside
Object handling: Locks	Able to handle locks
Object handling: Mail	Able to take the mail, carry it, deliver it.
Object handling: Milk	Able to open and close, milk carton, milk can, milk gallon, milk half gallon, includes carrying gallon of milk
Object handling: Nuts and bolts	Able to tighten nuts and bolts
Object handling: Opening pocket knife	Able to use a pocket knife
Object handling: Paper money	Able to hold paper money, fold it, stack it. Put it in and out of wallet/ purse.
Object handling: Peanut butter jar	Able to open a peanut butter jar
Object handling: Pegboard	Able to put pegs on a pegboard
Object handling: Pen	Able to use a pen for writing
Object handling: Pencil	Able to use a pencil for writing
Object handling: Picking golf balls from jar	Able to pick golf balls and put them in a jar, and take them out.
Object handling: Pictures/Photos	Able to take a picture using camera or phone
Object handling: Pills	Able to pick up pills, they drop
Object handling: Pins	Able to pick up pins, hair pins, clips
Object handling: Remote	Able to use a TV remote
Object handling: Rubber band	Able to use a rubber band to tie items, tie hair
Object handling: Scissors	Able to use scissors
Object handling: Screw - turning	Able to turn a screw with a screwdriver
Object handling: Screwdriver into screw	Able to place a screwdriver into a screw (before turning it)
Object handling: Seatbelt use	Able to put on the seatbelt
Object handling: Soap	Able to handle soap, slips off
Object handling: Socks	Able to put on socks
Object handling: Spoon	Able to use spoon, fork for eating, stirring.

Appendix D. Item Bank with Descriptions (Continued)

Item	***Description***
Object handling: Straw	Able to use a straw, place it and drink from it.
Object handling: Technology related: Computer related tasks	Able to do computer related tasks such as browsing the internet, finding information, Able to drag and drop objects on the screen.
Object handling: Technology related: iPad	Able to use the iPad, including apps on the device.
Object handling: Technology related: iPhone	Able to use the iPhone
Object handling: Technology related: Keypad use	Able to use the keypad for ATM, grocery store cashier, etc.
Object handling: Technology related: Mouse	Able to use the computer mouse
Object handling: Technology related: Phone	Able to use any phone
Object handling: Technology related: Typing	Able to type on a keyboard
Object handling: Technology related: Using a touch pad	Able to use a touch pad for a tablet computer or banking
Object handling: Technology related: Video games	Able to play video games
Object handling: Tray	Able to carry a tray
Object handling: Using a mobility device	Able to use a mobility device such as a cane or walker or crutches
Object handling: Wallet: Items in and Out	Able to transfer item s in and out of wallet
Object handling: Ziploc	Able to close and open a Ziploc® bag (press n seal type and sliding type)
Object handling: Zipper	Able to open and close a zipper
Other: Lighting a cigarette	Able to light a cigarette
Reach: Shelf height	Able to reach shelves at various heights high medium and low.
Reach: Take a box off a shelf	Able to get a box off a shelf
Recreation: Biking	Able to go biking using a bicycle or tricycle
Recreation: Gardening	Able to do gardening tasks, plant, weed, water, etc
Recreation: General	Able to engage in various recreational tasks like Swimming, watch TV, watch sports, fishing, playing with son, reading, paddling, woodworking, jewelry work, kayaking, quilting, photography, telegraph, singing, writing stories, geneology research, out in the nature, walking, camping, painting, watching sports, polish rocks, sewing
Recreation: Play music	Able to play music with an instrument
Recreation: Reading	Able to read, includes holding a book
Recreation: Sewing	Able to sew includes threading a needle, using a sewing machine
Recreation: Shuffling cards	Able to shuffle cards
Recreation: Wrapping a present	Able to wrap a present
Recreation: Yard work	Able to do yard work - mowing lawn
Self Care: Bath	Able to bathe includes rinsing the body, applying soap, drying the body.
Self Care: Brushing lint off clothes	Able to brush lint off clothes
Self Care: Buttoning	Able to button clothes
Self Care: Cleaning ears using Q-tip	Able to use Q tip to clean ears
Self Care: Dressing	Able to dress self includes putting on dress, shirt, pants, underclothes
Self Care: Dressing: Shoes and shoelaces	Able to don shoes and tie shoe laces

Appendix D. Item Bank with Descriptions (Continued)

Item	***Description***
Self Care: Dressing: Wearing bra	Able to wear a bra
Self Care: Eating: Cutting meat	Able to cut the meat on the plate for eating
Self Care: Eating: General	Able to bring spoon/ fork with food to mouth without spilling
Self Care: Eating: Sandwich	Able to eat a sandwich
Self Care: General	Able to take care of self
Self Care: Grooming: Nails - finger/toe	Able to take care of nails includes cutting, applying nail polish
Self Care: Grooming: Putting on make up/contact lens	Able to push on make up includes contact lens
Self Care: Grooming: Shaving	Able to shave hair on face and body
Self Care: Grooming: Wear Jewelry	Able to wear jewelry, putting on earring, fasten the back of the necklace.
Self Care: Hair care	Able to comb and tie hair
Self Care: Hygiene	Able to maintain general hygiene
Self Care: Oral care: Brushing teeth	Able to brush teeth
Self Care: Oral Care: Flossing	Able to floss teeth
Self Care: Oral care: Using toothpick	Able to use a toothpick
Self Care: Removing Band aid	Able to remove or apply band aid
Self Care: Scratch back	Able to scratch back
Self Care: Shower	Able to shower using a stand up shower
Self Care: Sleep issues	Issues with sleep due to arm and hand problems
Self Care: Toileting: Bowel and bladder care	Able to care for own bowel and bladder needs
Self Care: Toileting: General	Able to maintain toilet hygiene
Self Care: Toileting: Using handheld urinal	Able to use handheld urinal
Self Care: Toothpaste on toothbrush	Able to apply toothpaste on toothbrush
Self Care: Wash hands	Able to wash hands
Self Care: Washing face	Able to wash face
Self Care: Washing hair	Able to wash hair
Self Care: Wring out a washcloth	Able to wring out a washcloth
Transportation: Bus	Able to use bus for transportation
Transportation: Driving	Able to drive

Appendix E. Self-report Section of the Hand and Arm Function Measure Administered at Cognitive Interviews

HAND AND ARM FUNCTION QUESTIONNAIRE

Name____

Age________________Gender _____________Date________________________

Diagnosis(es)____

Current hand preference ☐ R ☐ L ☐ Both______________________________

Most affected hand ☐R ☐ L ☐ Both__________________________________

Least affected or unaffected hand: ☐ R ☐ L ☐Both ________________________

Do you have any muscle, nerve or skin issues with the hand and/or arm?

____

Medications taken today, dose and time_______________________________

____

Currently residing ☐Hospital ☐Home ☐Nursing Home ☐Senior Care ☐ Other

---

Instructions: In this questionnaire, we want to know how difficulties with hand and arm function may have interfered with your daily activities in the **past week**. Write N/A (Not Applicable) if you never have the symptom.

For the questions below, we want to know about your **RIGHT** hand.

SYMPTOMS: **RIGHT HAND**

#	Do these symptoms interfere with RIGHT hand function?	Not at all (3)	A Little (2)	Quite a lot (1)
1.	Coordination problems			
2.	Deformity			
3.	Fatigue			
4.	Hypersensitivity			
5.	Pain			
6.	Sensation loss			
7.	Slow movements			
8.	Spasticity			
9.	Stiffness			
10.	Tremor			
11.	Weakness			
			Subtotal	

For the questions below, we want to know about your **LEFT** hand.

SYMPTOMS: **LEFT HAND**

#	Do these symptoms interfere with LEFT hand function?	Not at all (3)	A Little (2)	Quite a lot (1)
12.	Coordination problems			
13.	Deformity			
14.	Fatigue			
15.	Hypersensitivity			
16.	Pain			
17.	Sensation loss			
18.	Slow movements			
19.	Spasticity			
20.	Stiffness			
21.	Tremor			
22.	Weakness			
			Subtotal	

For the questions below, think about using one or both of your hands.
Write NA (Not Applicable) if you never do the task.

Self care tasks

#	Do your hand function problems interfere with…	Not at all (3)	A Little (2)	Quite a lot (1)
23.	Brushing teeth			
24.	Feeding yourself (food to mouth)			
25.	Washing hair			
26.	Getting dressed			
27.	Putting on your shoes			
28.	Using clothing fasteners such as zipper or buttons			
29.	Putting on watch or jewelry			
30.	Toilet hygiene			
31.	Cutting your finger nails			
32.	Opening a childproof pill bottle			
33.	Sleeping			
			Subtotal	

For the questions below, think about using one or both of your hands.
Write NA (Not Applicable) if you never do the task.

Kitchen related tasks

#	Do your hand function problems interfere with…	Not at all (3)	A Little (2)	Quite a lot (1)
34.	Cutting vegetables			
35.	Lifting a cup with liquid in it			
36.	Lifting a gallon of milk			
37.	Lifting a grocery bag			
38.	Opening a can			
39.	Opening a milk carton			
40.	Opening and closing a jar			
41.	Opening or closing a Ziploc® bag			
			Subtotal	

Miscellaneous tasks

#	Do your hand function problems interfere with…	Not at all (3)	A Little (2)	Quite a lot (1)
42.	Clapping your hands			
43.	Getting items in and out of your wallet or purse			
44.	Handling credit cards or money			
45.	Holding an open book			
46.	Shaking hands when greeting someone			
47.	Turning door knobs			
48.	Using a key to open a door lock			
			Subtotal	

Technology related tasks

#	Do your hand function problems interfere with…	Not at all (3)	A Little (2)	Quite a lot (1)
49.	Pressing small buttons on a phone or remote control			
50.	Using a computer keyboard			
51.	Using a computer mouse			
52.	Using a touch screen on a computer, laptop, phone, or tablet			
			Subtotal	

For the questions below, think about using either or both of your hands.
Write NA (Not Applicable) if you never do the task.

(OPTIONAL) School, Recreation tasks

#	Do your hand function problems interfere with…	Not at all (3)	A Little (2)	Quite a lot (1)
53.	School in general			
54.	Please list one task at school involving your hands that is most difficult and rate it as the other items ______________________			
55.	Recreational or leisure in general			
56.	Please list one recreational or leisure task involving your hands that is most difficult and rate it as the other items ______________________			
57.	Childcare in general			
58.	Please list one childcare task involving your hands that is most difficult and rate it as the other items ______________________			
59.	Work in general (paid or volunteer work)			
60.	Please list one task at work involving your hands that is most difficult and rate it as the other items ______________________			
			Subtotal	

Thank you for completing this questionnaire!

Scoring for Self-report Section of Hand and Arm Function Measure

Tasks	**Score**
Right hand symptom subtotal (out of 33)	
Left hand symptom subtotal (out of 33)	
Self Care Subtotal (out of 33)	
Kitchen Subtotal (out of 24)	
Miscellaneous Subtotal (out of 21)	
Technology Subtotal (out of 12)	
(Optional) Childcare, school, work and recreational subtotal (out of 24)	
Total Raw Score (out of 180)	

In order to compute percent score:

- Use Method I, if none of the items were missed (or marked NA) by the patient.
- Use Method II, if there are missed (or marked NA) responses.

METHOD I

$\text{Final Percent Score} = \frac{\text{TotalRawScore}}{\text{MaxPointsPossible}} \text{ x } 100$	

Final Percent Score is out of 100 (higher is better)

METHOD II (Use only if there were missed or NA items in responses)

Step 1	Number of missed tasks	
Step 2	$\text{MissedPoints} = \text{MissedTasks x } 3$	
Step 3	$\text{MaxPossiblePoints} = 180 - \text{MissedPoints}$	
Step 4	$\text{Final Percent Score} = \frac{\text{TotalRawScore}}{\text{MaxPointsPossible}} \text{ x } 100$	

Final Percent Score is out of 100 (higher is better)

Final Percent score is the self reported hand function ability for an individual to engage in daily activities.

Appendix F. Performance-based Section of the Hand and Arm Function Measure Administered at Cognitive Interviews

ARM AND HAND FUNCTION PERFORMANCE MEASURE INSTRUCTIONS

Materials

- 15 American pennies
- 15 Dry black eye peas
- Bowl (18 oz. Glass/ ceramic cereal bowl 6"diameter at the rim)
- Can of beans with contents sealed (1 lb. 16 oz. e.g., Bush's red kidney beans)
- Coffee mug empty (12 oz., ceramic or glass)
- Empty plastic spice bottle (height 4" with screw top lid)
- Masking tape
- Metal teaspoon
- Padlock and key (1 1/8th inch wide)
- Pencil (sharpened Standard 2HB at least 6" long Eraser optional)
- Plastic jar with screw top lid (sealed, diameter 2.7"; height 4.7" e.g., Jiff peanut butter jar)
- Stopwatch
- Tape Measure
- Writing paper (writing pad 8 ½" x 11 ¾" white ruled ¼" between horizontal lines)

Sensory testing kit

- American penny
- Cotton ball
- Kidney bean (large, dry, red)
- Nut and bolt (nut hexagonal diameter ½"; bolt diameter 3/8")
- One pair adult tube socks (opaque, thick, large, 12" long, unused; e.g., American hospital socks)
- Paper clip (28 mm)
- Standard metal key (1.6" long)

Writing sample address

- Type the address below in font size 18 (large)
- Arial typeface (sans serif), lamination preferred.

Jane Doe
4567 NE Brown St.
Greenwood, WA
98213

Positioning

The patient or client is referred to as the 'individual' in this test. The individual is seated at a table with feet flat on the floor for assessment (Fig. 1). If seated in a wheelchair, the wheelchair should be locked and armrests removed, provided person has good trunk control. The individual should be seated close to the table so the forearms can rest on the table. Any deviations to this seating must be noted. For consistency in testing, repeated assessments should be conducted in the same seating position. If similar seating is not possible, consistent height of the seat and table should be maintained.

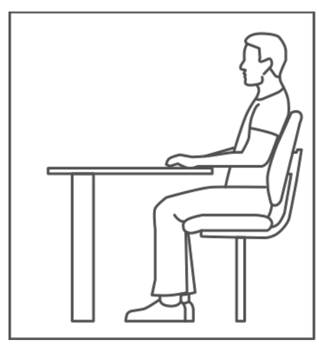


Fig. 1: Seating position

Reach distance and midline marking

Once the individual is seated at the testing table, you will need to determine reach distance and midline. To determine the reach distance, while the individual is seated upright, ask him/her to actively reach forward with both hands with the elbows fully extended, open hand, forearm pronated and flat on the table. Individual's back should stay in contact with the seatback during this measurement. The point where the wrist is located is considered the reach distance.

Mark the reach distance with a piece of masking tape on the table placed horizontally parallel to the front edge of the table. At the individual's midline, place masking tape extending from the reach distance marker to the front edge of the table. Measure and note the reach distance (Fig. 2). Use this distance for any repeated testing sessions. This is essential for evidence for reliability of scores. Place a 6-inch marker on either side of this midline marker.

Accommodations

If individual is unable achieve full elbow extension, the therapist can mark the reach distance of the wrist at their maximal active elbow extension and write a note indicating this modification. If an individual has severe limitations with one or both arms, use the less involved side for determining the reach distance.

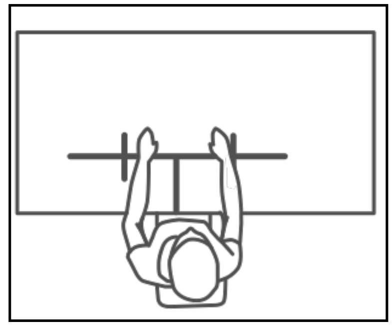


Fig 2: Reach distance, midline and 6 inch marking

Administration

Prior to testing, all the materials should be cleaned and ready to use. The instructions given under each task must be read aloud to the individual. Interpreters should be used for non-English speakers. Individuals being tested should be informed to wear their glasses during the assessment session if they routinely use them. The therapist should ensure that the person's sleeves are not restricting reach. For hygienic concerns, dry beans (black eyed peas and kidney beans for sensory test) should only be used one time and then discarded. If items such as bean cans and peanut butter jar break or leak, they should be replaced. For timed items, if objects are dropped on the floor, the individual must continue with the task while the therapist replaces the dropped item (e.g., putting a coin back in the bowl). However, if during a timed task a dropped item (e.g., bowl) causes the individual to completely stop the task, the task may be re-administered up to 3 times.

Demonstration

If therapist determines that individual is having difficulty understanding the item, tasks may be demonstrated.

Scoring Description

Score	Description	Details
5	Task completed independently	Task is completed within a reasonable time frame and with good quality of movement
4	Task completed independently with compensation	*Compensation* - 1) increased time 2) modified technique (e.g. using less affected/ intact side to assist, shoulder shrugs, trunk leaning) 3) adaptive devices or orthotics
3	Task completed with difficulty	*Difficulty* – Observed impairment related to 1) poor quality of movement 2) pain 3) extensive time 4) task compensations that compromise safety (i.e. dropped items)
2	Task partially completed	Independently completes $\geq$ 50% of task.
1	Task Attempted	Task attempted, less than 50% of task completed independently
0	Not able to do	Task not attempted, unable to carry out any part of the task

HAND AND ARM FUNCTION PERFORMANCE MEASURE

Name____

Age_______Gender_____________ Date____________________________________

Diagnosis(es) ____

Therapist Name ____

Assessment □**Initial** ☐**Follow up** ☐**Discharge** ______________________________

Most affected hand ☐ **R** ☐ **L** ☐**Both**____

Least or unaffected hand ☐ **R** ☐ **L** ☐**Both**__________________________________

Current Dominance ☐ **R** ☐ **L** ☐**Both** _______________________________________

Previous Dominance: ☐ **R** ☐ **L** ☐**Both** ☐**Same**_____________________________

Any muscle, nerve or skin issues with the hand and/or arm____

Medications taken today, dose and time____________________________________
____

Pain in arm or hand ____

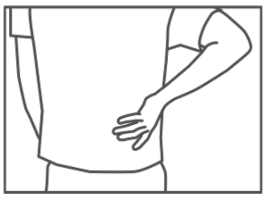
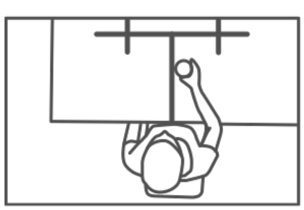
Right (no pain) 0 1 2 3 4 5 6 7 8 9 10 (severe)

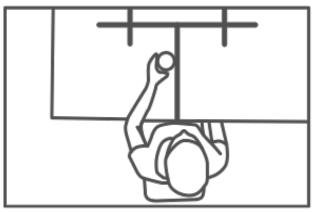
Left (no pain) 0 1 2 3 4 5 6 7 8 9 10 (severe)

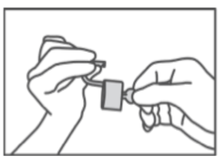
Pain location: **Right** ______________________**Left**__________________________

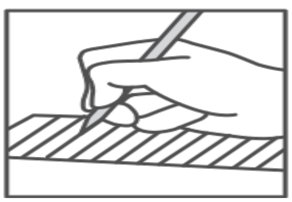
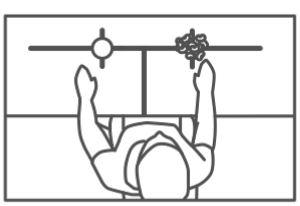
Adaptive devices or orthotics used during the test_____________________________
____

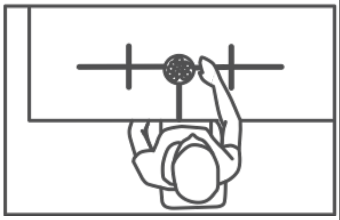
Therapist Instructions		**Individual Instructions**
Prepare all materials before starting the test. Administer each item and mark the task score. Give the individual as much time as needed to complete each task. As you administer the tasks, please write down any difficulties observed and any strategies the individual uses during the task. For example, compensation could be using both hands for unimanual tasks, leaning the trunk, bracing against the table, adaptive devices, compensations for vision issues such as moving objects closer or to the side, etc. Directions to be read to the individual are in italics. Information for the administrator is in regular type. The therapist may demonstrate the task to the individual along with the verbal instructions. The therapist may change the order of hands for unimanual tasks (left and then right) if desired.	Seating surface (e.g., chair): __________________________ Table specifications (e.g., height): __________________________ Reach distance (forward): __________________________ **Individual position:** Individual seated at the table with elbows on the table. Please keep the seating and table arrangement consistent for repeated testing.	*"For this test, you are asked to move your arm and hand, and reach, grasp and place some objects. This will help us learn how you use your hands for daily activities. Some tasks at the end of the test are timed and I will tell you when the timed tasks begin. I will show you how to do each task as we go along if needed. There are 10 tasks that I want you to do."* *"Do you have any questions? Are you ready to begin?"*

<u>Unimanual Tasks</u> **Instructions: *"In the first set of tasks, you will have to move the hand and arm on each side separately, right and then left."***	**SIDE**	**Indepen-dently** **(5)**	**With compen-sation** **(4)**	**With difficulty** **(3)**	**Partially completed** **(2)**	**Attempted** **(1)**	**Not able to do** **(0)**
1. Lower back 	Right						
"Take your right hand behind your body to touch your lower back with the palm of your hand." "Now use your left hand." Therapist Notes:	Left						
2. Coffee mug Set up: Therapist places an empty coffee mug at midline on the table at reach distance. ***"Pick up the coffee cup with your right hand, pretend to drink from it and place it back on the table."***	Right						
 **"Now use your left hand."** Preferred hand: __________ Therapist Notes:	Left						
3. Bean can RIGHT HAND Place unopened bean can on the 6inch marker on the RIGHT side. ***"Using your right hand, move the bean can back and forth 10 times. Be sure to stay on the 6 inch marker on either side and not drag the can."*** Therapist Notes:	Move LEFT						
	Move RIGHT						

Unimanual Tasks (Continued)	SIDE	Indepen-dently (5)	With compen-sation (4)	With difficulty (3)	Partially completed (2)	Attempted (1)	Not able to do (0)
4. Bean can LEFT HAND Set up: Place unopened bean can on the LEFT side of the midline marker starting at the LEFT reach distance 6 inches away. A 6-inch marker should be placed on either side of the midline. ***"Using your left hand, move the bean can back and forth 10 times. Be sure to stay on the 6 inch marker on either side and not drag the can."*** Therapist Notes: 	Move RIGHT						
	Move LEFT						
5. Getting jar Set up: Therapist holds a closed plastic jar just above individual's shoulder height at reach distance. ***"Grab this jar with your right hand and place it on the table."*** ***"Now use your left hand."*** Therapist notes: 	Right						
	Left						
Unimanual Subtotal (out of 80):							

Bimanual Tasks **Instructions: “*In the next set of tasks you need to use both hands together.*”**	**SIDE**	**Indepen-dently** **(5)**	**With compen-sation** **(4)**	**With difficulty** **(3)**	**Partially completed** **(2)**	**Attempted** **(1)**	**Not able to do** **(0)**
6. Opening jar Set up: Place an unopened  screw top plastic jar on the table at midline reach distance away. ***“Open this jar using both of your hands. Place the lid on the table when you are done.”*** Therapist Notes:	Hand on lid □ **R** □ **L** Mark one						
7. Open lock  Set up: Place a pad lock (in locked position) with key next to it on the table at the midline reach distance. ***“Open this lock with the key using both your hands.”*** Therapist Notes:	Key in hand □ **R** □ **L** Mark one						
8. Open spice bottle  Set up: Place an empty spice bottle (twist type in closed position) on the table. ***“Open this spice bottle using both your hands. Place the lid on the table when done.”*** Therapist Notes:	Hand on lid □ **R** □ **L** Mark one						
Bimanual Subtotal (out of 15):							

TIMED TASKS **Instructions:** ***"The last 4 tasks are timed. I will use a stopwatch to see how long it takes you. Try to do the task as quickly as you can."***	**SIDE**	**Indepen-dently** **(5)**	**With compen-sation** **(4)**	**With difficulty** **(3)**	**Partially completed** **(2)**	**Attempted** **(1)**	**Not able to do** **(0)**
9. Address writing Set up: Place paper and pencil at midline at reach distance on the table. Individual should use their current dominant hand for writing and place the other hand on the paper to stabilize the paper. Paper can be moved to get a good arm and hand position for writing without losing points. Printed address placed next to the individual. ***"Copy this address on the paper. Please use your glasses if you need for this task. Ready, set, go."*** Cursive or Print:___________________ *Therapist Notes:* 	TIME _____ Writing hand □ **R** □ **L** Mark one						
10. Peas in bottle Set up: Place the 15 black eyed peas on individuals current dominant side 6 inches from the midline and an open spice bottle on the individual's non-dominant side on the table 6 inches from midline. ***"Put the peas in the bottle one at a time. You can use your other hand to hold the bottle. Put them in as quickly as you can. Ready, set, go."***  *Therapist Notes:*	TIME _____ Peas in hand □ **R** □ **L** Mark one						

TIMED TASKS (Continued)	SIDE	Indepen-dently (5)	With compen-sation (4)	With difficulty (3)	Partially completed (2)	Attempted (1)	Not able to do (0)
11. Pennies on table. Set up: Place 15 pennies in a  bowl in front of the individual. Bowl placed within the reach distance at midline. ***"Pick up one penny at a time and place it on the table next to the bowl. Your other hand can steady the bowl. Ready, set go."*** *Therapist Notes:*	TIME ____ Pennies in hand □ **R** □ **L** Mark one						
12. Pennies in bowl – pick up and hold Set up: Place 5 pennies in a bowl. Bowl is placed within the reach distance at midline.  This is an in-hand manipulation (translation) task. ***"Pick up these 5 pennies one at a time. Keep holding the pennies in your hand as you pick the next one. Your other hand can steady the bowl. Ready, set, go"*** *Therapist Notes:*	TIME ____ Pennies in hand □ **R** □ **L** Mark one						
Timed Task Subtotal (out of 20):							

Scoring

Add the Unimanual, Bimanual and Timed subtotals to get the Total Raw Score.

Unimanual subtotal (out of 80)		Total Raw Score (out of 115)
Bimanual subtotal (out of 15)		
Timed task subtotal (out of 20)		

To convert Raw Score into Percent Score

$\text{Final Percent Score} = \frac{\text{TotalRawScore}}{115} \times 100$	Percent Score (out of 100)

Final Percent Score is out of 100 (higher is better) and it is the hand function ability for an individual to engage in daily tasks involving reach, grasp, manipulation, and dexterity.

REFERENCES

Abdullah, H. A., Tarry, C., Lambert, C., Barreca, S., & Allen, B. O. (2011). Results of clinicians using a therapeutic robotic system in an inpatient stroke rehabilitation unit. *Journal of Neuroengineering and Rehabilitation*, *8*, 50. http://doi.org/10.1186/1743-0003-8-50

Adamovich, S. V., Fluet, G. G., Mathai, A., Qiu, Q., Lewis, J., & Merians, A. S. (2009). Design of a complex virtual reality simulation to train finger motion for persons with hemiparesis: a proof of concept study. *Journal of Neuroengineering and Rehabilitation*, *6*, 28. http://doi.org/10.1186/1743-0003-6-28

Ahmed, S., Mayo, N. E., Higgins, J., Salbach, N. M., Finch, L., & Wood-Dauphinée, S. L. (2003). The Stroke Rehabilitation Assessment of Movement (STREAM): a comparison with other measures used to evaluate effects of stroke and rehabilitation. *Physical Therapy*, *83*(7), 617–630.

Alberta, F. G., ElAttrache, N. S., Bissell, S., Mohr, K., Browdy, J., Yocum, L., & Jobe, F. (2010). The development and validation of a functional assessment tool for the upper extremity in the overhead athlete. *The American Journal of Sports Medicine*, *38*(5), 903–911. http://doi.org/10.1177/0363546509355642

Allen, M. J., & Yen, W. M. (2002). *Introduction to measurement theory*. Prospect Heights, Ill.: Waveland Press.

Al Snih, S., Markides, K. S., Ottenbacher, K. J., & Raji, M. A. (2004). Hand grip strength and incident ADL disability in elderly Mexican Americans over a seven-year period. *Aging Clinical and Experimental Research*, *16*(6), 481–486.

Alt Murphy, M., Resteghini, C., Feys, P., & Lamers, I. (2015). An overview of systematic reviews on upper extremity outcome measures after stroke. *BMC Neurology*, *15*, 29. http://doi.org/10.1186/s12883-015-0292-6

Alt Murphy, M., Willén, C., & Sunnerhagen, K. S. (2012). Movement kinematics during a drinking task are associated with the activity capacity level after stroke. *Neurorehabilitation and Neural Repair*, *26*(9), 1106–1115. http://doi.org/10.1177/1545968312448234

American Educational Research Association. (1999). Validity: Part 1, Standards of Educational testing.

American Occupational Therapy Association. (2010). Workforce trends in Occupational therapy. American Occupational Therapy Association. Retrieved from https://www.aota.org/-/media/Corporate/Files/EducationCareers/Prospective/Workforce-trends-in-OT.PDF

Anderson, C. (2010). Presenting and Evaluating Qualitative Research. *American Journal of Pharmaceutical Education*, *74*(8), 1–7.

Angst, F., Goldhahn, J., Drerup, S., Kolling, C., Aeschlimann, A., Simmen, B. R., & Schwyzer, H. K. (2012). Responsiveness of five outcome measurement instruments in total elbow arthroplasty. *Arthritis Care & Research*, *64*(11), 1749–1755. http://doi.org/10.1002/acr.21744

Annear, M. J., Toye, C. M., Eccleston, C. E., McInerney, F. J., Elliott, K. E. J., Tranter, B. K., … Robinson, A. L. (2015). Dementia Knowledge Assessment Scale: Development and Preliminary Psychometric Properties. *Journal of the American Geriatrics Society*, *63*(11), 2375–2381. http://doi.org/10.1111/jgs.13707

Arya, K. N., & Pandian, S. (2013). Effect of task-based mirror therapy on motor recovery of the upper extremity in chronic stroke patients: a pilot study. *Topics in Stroke Rehabilitation*, *20*(3), 210–217. http://doi.org/10.1310/tsr2003-210

Asher, I. (2007). *Occupational Therapy Assessment Tools: An Annotated Index, 3rd Edition* (3rd ed.). AOTA Press.

Ashford, S., Slade, M., Malaprade, F., & Turner-Stokes, L. (2008). Evaluation of functional outcome measures for the hemiparetic upper limb: a systematic review. *Journal of Rehabilitation Medicine: Official Journal of the UEMS European Board of Physical and Rehabilitation Medicine*, *40*(10), 787–795. http://doi.org/10.2340/16501977-0276

Ashford, S., & Turner-Stokes, L. (2013). Systematic review of upper-limb function measurement methods in botulinum toxin intervention for focal spasticity. *Physiotherapy Research International: The Journal for Researchers and Clinicians in Physical Therapy*, *18*(3), 178–189. http://doi.org/10.1002/pri.1554

Ashford, S., Turner-Stokes, L., Siegert, R., & Slade, M. (2013). Initial psychometric evaluation of the Arm Activity Measure (ArmA): a measure of activity in the hemiparetic arm. *Clinical Rehabilitation*, *27*(8), 728–740. http://doi.org/10.1177/0269215512474942

Au-Yeung, S. S. Y., & Hui-Chan, C. W. Y. (2014). Electrical acupoint stimulation of the affected arm in acute stroke: a placebo-controlled randomized clinical trial. *Clinical Rehabilitation*, *28*(2), 149–158. http://doi.org/10.1177/0269215513494875

Baker, K., Barrett, L., Playford, E. D., Aspden, T., Riazi, A., & Hobart, J. (2015). Measuring arm function early after stroke: is the DASH good enough? *Journal of Neurology, Neurosurgery, and Psychiatry*. http://doi.org/10.1136/jnnp-2015-310557

Barreca, S., Gowland, C., Stratford, P., Huijbregts, M., Griffiths, J., Torresin, W., … Masters, L. (2004). Development of the Chedoke Arm and Hand Activity Inventory: theoretical constructs, item generation, and selection. *Topics in Stroke Rehabilitation*, *11*(4), 31–42. http://doi.org/10.1310/JU8P-UVK6-68VW-CF3W

Barreca, S. R., Stratford, P. W., Lambert, C. L., Masters, L. M., & Streiner, D. L. (2005). Test-retest reliability, validity, and sensitivity of the Chedoke arm and hand activity inventory: a new measure of upper-limb function for survivors of stroke. *Archives of Physical Medicine and Rehabilitation*, *86*(8), 1616–1622. http://doi.org/10.1016/j.apmr.2005.03.017

Barreca, S. R., Stratford, P. W., Masters, L. M., Lambert, C. L., Griffiths, J., & McBay, C. (2006). Validation of Three Shortened Versions of the Chedoke Arm and Hand Activity Inventory. *Physiotherapy Canada*, *58*(2), 148–156. http://doi.org/10.3138/ptc.58.2.148

Bartolo, M., De Nunzio, A. M., Sebastiano, F., Spicciato, F., Tortola, P., Nilsson, J., & Pierelli, F. (2014). Arm weight support training improves functional motor outcome and movement smoothness after stroke. *Functional Neurology*, *29*(1), 15–21.

Beebe, J. A., & Lang, C. E. (2009). Relationships and responsiveness of six upper extremity function tests during the first six months of recovery after stroke. *Journal of Neurologic Physical Therapy: JNPT*, *33*(2), 96–103. http://doi.org/10.1097/NPT.0b013e3181a33638

Berglund, K., & Fugl-Meyer, A. R. (1986). Upper extremity function in hemiplegia. A cross-validation study of two assessment methods. *Scandinavian Journal of Rehabilitation Medicine*, *18*(4), 155–157.

Bernspång, B., Asplund, K., Eriksson, S., & Fugl-Meyer, A. R. (1987). Motor and perceptual impairments in acute stroke patients: effects on self-care ability. *Stroke; a Journal of Cerebral Circulation*, *18*(6), 1081–1086.

Blennerhassett, J. M., Avery, R. M., & Carey, L. M. (2010). The test-retest reliability and responsiveness to change for the Hand Function Survey during stroke rehabilitation. *Australian Occupational Therapy Journal*, *57*(6), 431–438. http://doi.org/10.1111/j.1440-1630.2010.00884.x

Bohannon, R. (1999). Motricity Index scores are valid indicator of paretic upper extremity strength following stroke. *Journal of Physical Therapy Science*, *11*, 59–61.

Borstad, A. L., Bird, T., Choi, S., Goodman, L., Schmalbrock, P., & Nichols-Larsen, D. S. (2013). Sensorimotor training and neural reorganization after stroke: a case series. *Journal of Neurologic Physical Therapy: JNPT*, *37*(1), 27–36. http://doi.org/10.1097/NPT.0b013e318283de0d

Bowman, M. H., Taub, E., Uswatte, G., Delgado, A., Bryson, C., Morris, D. M., … Mark, V. W. (2006). A treatment for a chronic stroke patient with a plegic hand combining CI therapy with conventional rehabilitation procedures: case report. *NeuroRehabilitation*, *21*(2), 167–176.

Boyne, P., Dunning, K., Levine, P., Hermann, V. H., & Page, S. J. (2010). Orthotic functional electrical stimulation following botulinum toxin for a young adult with severe hand impairment due to childhood stroke. *Physiotherapy Theory and Practice*, *26*(4), 267–274. http://doi.org/10.3109/09593980902967188

Brauer, S. G., Hayward, K. S., Carson, R. G., Cresswell, A. G., & Barker, R. N. (2013). The efficacy of SMART Arm training early after stroke for stroke survivors with severe upper limb disability: a protocol for a randomised controlled trial. *BMC Neurology*, *13*, 71. http://doi.org/10.1186/1471-2377-13-71

Broderick, J. E., Schneider, S., Junghaenel, D. U., Schwartz, J. E., & Stone, A. A. (2013). Validity and reliability of patient-reported outcomes measurement information system instruments in osteoarthritis. *Arthritis Care & Research*, *65*(10), 1625–1633. http://doi.org/10.1002/acr.22025

Broeks, J. G., Lankhorst, G. J., Rumping, K., & Prevo, A. J. (1999). The long-term outcome of arm function after stroke: results of a follow-up study. *Disability and Rehabilitation*, *21*(8), 357–364.

Brogårdh, C., & Lexell, J. (2010). A 1-year follow-up after shortened constraint-induced movement therapy with and without mitt poststroke. *Archives of Physical Medicine and Rehabilitation*, *91*(3), 460–464. http://doi.org/10.1016/j.apmr.2009.11.009

Brogårdh, C., & Sjölund, B. H. (2006). Constraint-induced movement therapy in patients with stroke: a pilot study on effects of small group training and of extended mitt use. *Clinical Rehabilitation*, *20*(3), 218–227.

Brunner, I. C., Skouen, J. S., & Strand, L. I. (2012). Is modified constraint-induced movement therapy more effective than bimanual training in improving arm motor function in the subacute phase post stroke? A randomized controlled trial. *Clinical Rehabilitation*, *26*(12), 1078–1086. http://doi.org/10.1177/0269215512443138

Brunnstrom, S. (1966). Motor testing procedures in hemiplegia: based on sequential recovery stages. *Physical Therapy*, *46*(4), 357–75.

Buneo, C. A., Jarvis, M. R., Batista, A. P., & Andersen, R. A. (2002). Direct visuomotor transformations for reaching. *Nature*, *416*(6881), 632–636. http://doi.org/10.1038/416632a

Burns, A., Burridge, J., & Pickering, R. (2007). Does the use of a constraint mitten to encourage use of the hemiplegic upper limb improve arm function in adults with subacute stroke? *Clinical Rehabilitation*, *21*(10), 895–904. http://doi.org/10.1177/0269215507079144

Canny, M. L., Thompson, J. M., & Wheeler, M. J. (2009). Reliability of the box and block test of manual dexterity for use with patients with fibromyalgia. *The American Journal of Occupational Therapy: Official Publication of the American Occupational Therapy Association*, *63*(4), 506–510.

Cano, S. J., Barrett, L. E., Zajicek, J. P., & Hobart, J. C. (2011). Beyond the reach of traditional analyses: using Rasch to evaluate the DASH in people with multiple sclerosis. *Multiple Sclerosis (Houndmills, Basingstoke, England)*, *17*(2), 214–222. http://doi.org/10.1177/1352458510385269

Carlozzi, N. E., Tulsky, D. S., & Kisala, P. A. (2011). Traumatic brain injury patient-reported outcome measure: identification of health-related quality-of-life issues relevant to individuals with traumatic brain injury. *Archives of Physical Medicine and Rehabilitation*, *92*(10 Suppl), S52–60. http://doi.org/10.1016/j.apmr.2010.12.046

Carmeli, E., Patish, H., & Coleman, R. (2003). The aging hand. *The Journals of Gerontology. Series A, Biological Sciences and Medical Sciences*, *58*(2), 146–152.

Carod-Artal, F. J., Ferreira Coral, L., Stieven Trizotto, D., & Menezes Moreira, C. (2009). Self- and proxy-report agreement on the Stroke Impact Scale. *Stroke; a Journal of Cerebral Circulation*, *40*(10), 3308–3314. http://doi.org/10.1161/STROKEAHA.109.558031

Carr, J. H., Shepherd, R. B., Nordholm, L., & Lynne, D. (1985). Investigation of a new motor assessment scale for stroke patients. *Physical Therapy*, *65*(2), 175–180.

Carroll, D. (1965). A Quantitative test of upper extremity function. *Journal of Chronic Diseases*, *18*, 479–491.

Celik, O., O'Malley, M. K., Boake, C., Levin, H. S., Yozbatiran, N., & Reistetter, T. A. (2010). Normalized movement quality measures for therapeutic robots strongly correlate with clinical motor impairment measures. *IEEE Transactions on Neural Systems and Rehabilitation Engineering: A Publication of the IEEE Engineering in Medicine and Biology Society*, *18*(4), 433–444. http://doi.org/10.1109/TNSRE.2010.2047600

Cella, D. (2015). Neuro-QoL. Retrieved October 22, 2015, from http://www.neuroqol.org/Pages/default.aspx

Cella, D., Lai, J. S., Nowinski, C. J., Victorson, D., Peterman, A., Miller, D., … Moy, C. (2012). Neuro-QOL: brief measures of health-related quality of life for clinical research in neurology. *Neurology*, *78*(23), 1860–1867. http://doi.org/10.1212/WNL.0b013e318258f744

Chae, J., Harley, M. Y., Hisel, T. Z., Corrigan, C. M., Demchak, J. A., Wong, Y. T., & Fang, Z. P. (2009). Intramuscular electrical stimulation for upper limb recovery in chronic hemiparesis: an exploratory randomized clinical trial. *Neurorehabilitation and Neural Repair*, *23*(6), 569–578. http://doi.org/10.1177/1545968308328729

Chae, J., Labatia, I., & Yang, G. (2003). Upper limb motor function in hemiparesis: concurrent validity of the Arm Motor Ability test. *American Journal of Physical Medicine & Rehabilitation / Association of Academic Physiatrists*, *82*(1), 1–8. http://doi.org/10.1097/01.PHM.0000034950.95737.91

Chan, M. K. L., Tong, R. K. Y., & Chung, K. Y. K. (2009). Bilateral upper limb training with functional electric stimulation in patients with chronic stroke. *Neurorehabilitation and Neural Repair*, *23*(4), 357–365. http://doi.org/10.1177/1545968308326428

Chen, & Bode, R. K. (2010). Psychometric validation of the Manual Ability Measure-36 (MAM-36) in patients with neurologic and musculoskeletal disorders. *Archives of Physical Medicine and Rehabilitation*, *91*(3), 414–420. http://doi.org/10.1016/j.apmr.2009.11.012

Chen, Granger, C. V., Peimer, C. A., Moy, O. J., & Wald, S. (2005). Manual Ability Measure (MAM-16): a preliminary report on a new patient-centred and task-oriented outcome measure of hand function. *Journal of Hand Surgery (Edinburgh, Scotland)*, *30*(2), 207–216. http://doi.org/10.1016/j.jhsb.2004.12.005

Chen, H., Lin, K., Wu, C., & Chen, C. (2012). Rasch validation and predictive validity of the action research arm test in patients receiving stroke rehabilitation. *Archives of Physical Medicine and Rehabilitation*, *93*(6), 1039–1045. http://doi.org/10.1016/j.apmr.2011.11.033

Chen, H.M., Chen, C. C., Hsueh, I.P., Huang, S.L., & Hsieh, C.L. (2009). Test-retest reproducibility and smallest real difference of 5 hand function tests in patients with stroke. *Neurorehabilitation and Neural Repair*, *23*(5), 435–440. http://doi.org/10.1177/1545968308331146

Chen, H.M., Hsieh, C.L., Sing Kai Lo, Liaw, L.J., Chen, S.M., & Lin, J.H. (2007). The test-retest reliability of 2 mobility performance tests in patients with chronic stroke. *Neurorehabilitation and Neural Repair*, *21*(4), 347–352. http://doi.org/10.1177/1545968306297864

Chen, S., Lewthwaite, R., Schweighofer, N., & Winstein, C. J. (2013). Discriminant validity of a new measure of self-efficacy for reaching movements after stroke-induced hemiparesis. *Journal of Hand Therapy: Official Journal of the American Society of Hand Therapists*, *26*(2), 116–122; quiz 123. http://doi.org/10.1016/j.jht.2012.09.002

Chuang, L., Wu, C., & Lin, K. (2012). Reliability, validity, and responsiveness of myotonometric measurement of muscle tone, elasticity, and stiffness in patients with stroke. *Archives of Physical Medicine and Rehabilitation*, *93*(3), 532–540. http://doi.org/10.1016/j.apmr.2011.09.014

Cohen, J. A., Fischer, J. S., Bolibrush, D. M., Jak, A. J., Kniker, J. E., Mertz, L. A., … Cutter, G. R. (2000). Intrarater and interrater reliability of the MS functional composite outcome measure. *Neurology*, *54*(4), 802–806.

Collen, F. M., Wade, D. T., & Bradshaw, C. M. (1990). Mobility after stroke: reliability of measures of impairment and disability. *International Disability Studies*, *12*(1), 6–9.

Collin, C., & Wade, D. (1990). Assessing motor impairment after stroke: a pilot reliability study. *Journal of Neurology, Neurosurgery, and Psychiatry*, *53*(7), 576–579.

Combs, S. A., Kelly, S. P., Barton, R., Ivaska, M., & Nowak, K. (2010). Effects of an intensive, task-specific rehabilitation program for individuals with chronic stroke: a case series. *Disability and Rehabilitation*, *32*(8), 669–678. http://doi.org/10.3109/09638280903242716

Cook, D., & Beckman, T. (2006). Current concepts in validity and reliability for psychometric instruments: theory and application. *The American Journal of Medicine*, *119*(2), 166.e7–16.

Cook, K., Teal, C., Bjorner, J., Cella, D., Chang, C. H., Crane, P., … Reeve, B. (2007). IRT health outcomes data analysis project: an overview and summary. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, *16 Suppl 1*, 121–132. http://doi.org/10.1007/s11136-007-9177-5

Cronbach, L., & Meehl, P. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*(4), 281–302.

Crow, J. L., & Harmeling-van der Wel, B. C. (2008). Hierarchical properties of the motor function sections of the Fugl-Meyer assessment scale for people after stroke: a retrospective study. *Physical Therapy*, *88*(12), 1554–1567. http://doi.org/10.2522/ptj.20070186

Cutkosky, M., & Wright, P. (1986). Modeling Manufacturing Grips and Correlation with the Design of Robotic Hands. *Proceedings of IEEE International Conference on Robotics and Automation*, 1533–39.

Dahl, A. E., Askim, T., Stock, R., Langørgen, E., Lydersen, S., & Indredavik, B. (2008). Short- and long-term outcome of constraint-induced movement therapy after stroke: a randomized controlled feasibility trial. *Clinical Rehabilitation*, *22*(5), 436–447. http://doi.org/10.1177/0269215507084581

Dalbeth, N., Collis, J., Gregory, K., Clark, B., Robinson, E., & McQueen, F. M. (2007). Tophaceous joint disease strongly predicts hand function in patients with gout. *Rheumatology*, *46*(12), 1804–1807. http://doi.org/10.1093/rheumatology/kem246

Daley, K., Mayo, N., & Wood-Dauphinée, S. (1999). Reliability of scores on the Stroke Rehabilitation Assessment of Movement (STREAM) measure. *Physical Therapy*, *79*(1), 8–19; quiz 20–23.

Daly, J. J., Hogan, N., Perepezko, E. M., Krebs, H. I., Rogers, J. M., Goyal, K. S., … Ruff, R. L. (2005). Response to upper-limb robotics and functional neuromuscular stimulation following stroke. *Journal of Rehabilitation Research and Development*, *42*(6), 723–736.

Daly, J. J., & Ruff, R. L. (2007). Construction of efficacious gait and upper limb functional interventions based on brain plasticity evidence and model-based measures for stroke patients. *TheScientificWorldJournal*, *7*, 2031–2045. http://doi.org/10.1100/tsw.2007.299

Dean, C., & Mackey, F. (1992). Motor assessment scale scores as a measure of rehabilitation outcome following stroke. *Australian Journal of Physiotherapy*, *38*, 31–35.

Demeurisse, G., Demol, O., & Robaye, E. (1979). Motor evaluation in vascular hemiplegia. *Bruxelles Médical*, *59*(2), 87–94.

Demeurisse, G., Demol, O., & Robaye, E. (1980). Motor evaluation in vascular hemiplegia. *European Neurology*, *19*(6), 382–389.

Desrosiers, J., Bravo, G., Hébert, R., & Dutil, E. (1995). Normative data for grip strength of elderly men and women. *The American Journal of Occupational Therapy: Official Publication of the American Occupational Therapy Association*, *49*(7), 637–644.

Desrosiers, J., Hébert, R., Bravo, G., & Dutil, E. (1995). Upper extremity performance test for the elderly (TEMPA): normative data and correlates with sensorimotor parameters. Test d’Evaluation des Membres Supérieurs de Personnes Agées. *Archives of Physical Medicine and Rehabilitation*, *76*(12), 1125–1129.

Dettmann, M. A., Linder, M. T., & Sepic, S. B. (1987). Relationships among walking performance, postural stability, and functional assessments of the hemiplegic patient. *American Journal of Physical Medicine*, *66*(2), 77–90.

DeVellis, R. (2012). *Scale Development* (3rd ed.). Chapel Hill, North Carolina: SAGE Publications Ltd.

Devraj, R., & Wallace, L. S. (2013). Application of the content expert process to develop a clinically useful low-literacy Chronic Kidney Disease Self-Management Knowledge Tool (CKD-SMKT). *Research in Social & Administrative Pharmacy: RSAP*, *9*(5), 633–639. http://doi.org/10.1016/j.sapharm.2012.09.006

Doyle, P. J., McNeil, M. R., Bost, J. E., Ross, K. B., Wambaugh, J. L., Hula, W. D., & Mikolic, J. M. (2007). The Burden of Stroke Scale (BOSS) provided valid, reliable, and responsive score estimates of functioning and well-being during the first year of recovery from stroke. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, *16*(8), 1389–1398. http://doi.org/10.1007/s11136-007-9247-8

Drussell, R. D. (1959). Relationship of Minnesota Rate of Manipulation Test with the industrial work performance of the adult cerebral palsied. *The American Journal of Occupational Therapy: Official Publication of the American Occupational Therapy Association*, *13*(2, Part 2), 93–96 passim.

Duncan, P. W., Goldstein, L. B., Matchar, D., Divine, G. W., & Feussner, J. (1992). Measurement of motor recovery after stroke. Outcome assessment and sample size requirements. *Stroke; a Journal of Cerebral Circulation*, *23*(8), 1084–1089.

Duncan, P. W., Lai, S., Tyler, D., Perera, S., Reker, D., & Studenski, S. (2002). Evaluation of proxy responses to the Stroke Impact Scale. *Stroke; a Journal of Cerebral Circulation*, *33*(11), 2593–2599.

Duncan, P. W., Propst, M., & Nelson, S. G. (1983). Reliability of the Fugl-Meyer assessment of sensorimotor recovery following cerebrovascular accident. *Physical Therapy*, *63*(10), 1606–1610.

Duncan, P. W., Wallace, D., Lai, S. M., Johnson, D., Embretson, S., & Laster, L. J. (1999). The stroke impact scale version 2.0. Evaluation of reliability, validity, and sensitivity to change. *Stroke; a Journal of Cerebral Circulation*, *30*(10), 2131–2140.

Earhart, G. M., Cavanaugh, J. T., Ellis, T., Ford, M. P., Foreman, K. B., & Dibble, L. (2011). The 9-hole PEG test of upper extremity function: average values, test-retest reliability, and factors contributing to performance in people with Parkinson disease. *Journal of Neurologic Physical Therapy: JNPT*, *35*(4), 157–163. http://doi.org/10.1097/NPT.0b013e318235da08

Endres, M., Nyary, I., Banhidi, M., & Deak, G. (1990). Stroke rehabilitation: A method and evaluation. *International Journal of Rehabilitation Research. Internationale Zeitschrift Für Rehabilitationsforschung. Revue Internationale De Recherches De Réadaptation*, *13*, 225–236.

English, C. K., Hillier, S. L., Stiller, K., & Warden-Flood, A. (2006). The sensitivity of three commonly used outcome measures to detect change amongst patients receiving inpatient rehabilitation following stroke. *Clinical Rehabilitation*, *20*(1), 52–55.

Erasmus, L. P., Sarno, S., Albrecht, H., Schwecht, M., Pöllmann, W., & König, N. (2001). Measurement of ataxic symptoms with a graphic tablet: standard values in controls and validity in Multiple Sclerosis patients. *Journal of Neuroscience Methods*, *108*(1), 25–37.

Erkinjuntti, T., Sulkava, R., Wikström, J., & Autio, L. (1987). Short Portable Mental Status Questionnaire as a screening test for dementia and delirium among the elderly. *Journal of the American Geriatrics Society*, *35*(5), 412–416.

Ferreiro, K., dos Santos, R., & Conforto, A. (2010). Psychometric properties of the Portuguese version of the Jebsen-Taylor test for adults with mild hemiparesis. *Revista Brasileira de Fisioterapia (Brazilian Journal of Physiotherapy)*, *14*(5), 377–81.

Feys, P., Duportail, M., Kos, D., Van Asch, P., & Ketelaer, P. (2002). Validity of the TEMPA for the measurement of upper limb function in multiple sclerosis. *Clinical Rehabilitation*, *16*(2), 166–173.

Feys, P., Romberg, A., Ruutiainen, J., & Ketelaer, P. (2004). Interference of Upper Limb Tremor on Daily Life Activities in People with Multiple Sclerosis. *Occupational Therapy in Health Care*, *17*(3-4), 81–95. http://doi.org/10.1080/J003v17n03_06

Finch, E., Brooks, D., Stratford, P., & Mayo, N. (2002). *Physical Outcome Measures: A guide to enhance physical outcome meaures.* Ontario, Canada: Lippincott, Williams & Wilkins.

Fong, K., Ng, B., Chan, D., Chan, E., Ma, D., Au, B., … Chan, V. (2004). Development of the Hong Kong Version of the Functional Test for the Hemiplegic Upper Extremity (FTHUE-HK). *Hong Kong Journal of Occupational Therapy*, *14*(1), 21–29. http://doi.org/10.1016/S1569-1861(09)70025-7

Freitas, S., Simões, M. R., Marôco, J., Alves, L., & Santana, I. (2012). Construct Validity of the Montreal Cognitive Assessment (MoCA). *Journal of the International Neuropsychological Society: JINS*, *18*(2), 242–250.

Fritz, S. L., Blanton, S., Uswatte, G., Taub, E., & Wolf, S. L. (2009). Minimal detectable change scores for the Wolf Motor Function Test. *Neurorehabilitation and Neural Repair*, *23*(7), 662–667. http://doi.org/10.1177/1545968309335975

Gershon, R. C., Lai, J. S., Bode, R., Choi, S., Moy, C., Bleck, T., … Cella, D. (2012). Neuro-QOL: quality of life item banks for adults with neurological disorders: item development and calibrations based upon clinical and general population testing. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, *21*(3), 475–486. http://doi.org/10.1007/s11136-011-9958-8

Gershon, R. C., Rothrock, N., Hanrahan, R., Bass, M., & Cella, D. (2010). The use of PROMIS and assessment center to deliver patient-reported outcome measures in clinical research. *Journal of Applied Measurement*, *11*(3), 304–314.

Gladstone, D. J., Danells, C. J., & Black, S. E. (2002). The fugl-meyer assessment of motor recovery after stroke: a critical review of its measurement properties. *Neurorehabilitation and Neural Repair*, *16*(3), 232–240.

Golafshani, N. (2003). Understanding Reliability and Validity in Qualitative Research. *The Qualitative Report*, *8*(4), 597–607.

Gummesson, C., Atroshi, I., & Ekdahl, C. (2003). The disabilities of the arm, shoulder and hand (DASH) outcome questionnaire: longitudinal construct validity and measuring self-rated health change after surgery. *BMC Musculoskeletal Disorders*, *4*, 11. http://doi.org/10.1186/1471-2474-4-11

Gurcay, E., Bal, A., & Cakci, A. (2009). Health-related quality of life in first-ever stroke patients. *Annals of Saudi Medicine*, *29*(1), 36–40.

Gustafsson, L. A., Turpin, M. J., & Dorman, C. M. (2010). Clinical utility of the Chedoke Arm and Hand Activity Inventory for stroke rehabilitation. *Canadian Journal of Occupational Therapy. Revue Canadienne D'ergothérapie*, *77*(3), 167–173.

Hackel, M. E., Wolfe, G. A., Bang, S. M., & Canfield, J. S. (1992). Changes in hand function in the aging adult as determined by the Jebsen Test of Hand Function. *Physical Therapy*, *72*(5), 373–377.

Hammer, A. M., & Lindmark, B. (2010). Responsiveness and validity of the Motor Activity Log in patients during the subacute phase after stroke. *Disability and Rehabilitation*, *32*(14), 1184–1193. http://doi.org/10.3109/09638280903437253

Harris, J. E., Eng, J. J., Miller, W. C., & Dawson, A. S. (2009). A self-administered Graded Repetitive Arm Supplementary Program (GRASP) improves arm function during inpatient stroke rehabilitation: a multi-site randomized controlled trial. *Stroke; a Journal of Cerebral Circulation*, *40*(6), 2123–2128. http://doi.org/10.1161/STROKEAHA.108.544585

Hays, R. D., Spritzer, K. L., Amtmann, D., Lai, J.S., Dewitt, E. M., Rothrock, N., … Krishnan, E. (2013). Upper-extremity and mobility subdomains from the Patient-Reported Outcomes Measurement Information System (PROMIS) adult physical functioning item bank. *Archives of Physical Medicine and Rehabilitation*, *94*(11), 2291–2296. http://doi.org/10.1016/j.apmr.2013.05.014

Heller, A., Wade, D. T., Wood, V. A., Sunderland, A., Hewer, R. L., & Ward, E. (1987). Arm function after stroke: measurement and recovery over the first three months. *Journal of Neurology, Neurosurgery, and Psychiatry*, *50*(6), 714–719.

Hesse, S., Werner, C., Pohl, M., Rueckriem, S., Mehrholz, J., & Lingnau, M. L. (2005). Computerized arm training improves the motor control of the severely affected arm after stroke: a single-blinded randomized trial in two centers. *Stroke; a Journal of Cerebral Circulation*, *36*(9), 1960–1966. http://doi.org/10.1161/01.STR.0000177865.37334.ce

Higgins, J., Mayo, N. E., Desrosiers, J., Salbach, N. M., & Ahmed, S. (2005). Upper-limb function and recovery in the acute phase poststroke. *The Journal of Rehabilitation Research and Development*, *42*(1), 65. http://doi.org/10.1682/JRRD.2003.10.0156

Higgins, J., Salbach, N. M., Wood-Dauphinee, S., Richards, C. L., Côté, R., & Mayo, N. E. (2006). The effect of a task-oriented intervention on arm function in people with stroke: a randomized controlled trial. *Clinical Rehabilitation*, *20*(4), 296–310.

Hijmans, J. M., Hale, L. A., Satherley, J. A., McMillan, N. J., & King, M. J. (2011). Bilateral upper-limb rehabilitation after stroke using a movement-based game controller. *Journal of Rehabilitation Research and Development*, *48*(8), 1005–1013.

Hsieh, Y., Wu, C., Lin, K., Chang, Y., Chen, C., & Liu, J. (2009). Responsiveness and validity of three outcome measures of motor function after stroke rehabilitation. *Stroke; a Journal of Cerebral Circulation*, *40*(4), 1386–1391. http://doi.org/10.1161/STROKEAHA.108.530584

Hsueh, I.P., Hsu, M.J., Sheu, C.F., Lee, S., Hsieh, C.L., & Lin, J.H. (2008). Psychometric comparisons of 2 versions of the Fugl-Meyer Motor Scale and 2 versions of the Stroke Rehabilitation Assessment of Movement. *Neurorehabilitation and Neural Repair*, *22*(6), 737–744. http://doi.org/10.1177/1545968308315999

Hsueh, I.P., Wang, C.H., Sheu, C.F., & Hsieh, C.L. (2003). Comparison of psychometric properties of three mobility measures for patients with stroke. *Stroke; a Journal of*

Cerebral Circulation, *34*(7), 1741–1745. http://doi.org/10.1161/01.STR.0000075295.45185.D4

Hsu, J. E., Nacke, E., Park, M. J., Sennett, B. J., & Huffman, G. R. (2010). The Disabilities of the Arm, Shoulder, and Hand questionnaire in intercollegiate athletes: validity limited by ceiling effect. *Journal of Shoulder and Elbow Surgery / American Shoulder and Elbow Surgeons ... [et Al.]*, *19*(3), 349–354. http://doi.org/10.1016/j.jse.2009.11.006

Huang, Y., Wu, C., Hsieh, Y., & Lin, K. (2010). Predictors of change in quality of life after distributed constraint-induced therapy in patients with chronic stroke. *Neurorehabilitation and Neural Repair*, *24*(6), 559–566. http://doi.org/10.1177/1545968309358074

Hudak, P. L., Amadio, P. C., & Bombardier, C. (1996). Development of an upper extremity outcome measure: the DASH (disabilities of the arm, shoulder and hand) [corrected]. The Upper Extremity Collaborative Group (UECG). *American Journal of Industrial Medicine*, *29*(6), 602–608. http://doi.org/10.1002/(SICI)1097-0274(199606)29:6<602::AID-AJIM4>3.0.CO;2-L

Hughes, A. M., Freeman, C. T., Burridge, J. H., Chappell, P. H., Lewin, P. L., & Rogers, E. (2009). Feasibility of iterative learning control mediated by functional electrical stimulation for reaching after stroke. *Neurorehabilitation and Neural Repair*, *23*(6), 559–568. http://doi.org/10.1177/1545968308328718

Huisstede, B. M. A., Feleus, A., Bierma-Zeinstra, S. M., Verhaar, J. A., & Koes, B. W. (2009). Is the disability of arm, shoulder, and hand questionnaire (DASH) also valid and responsive in patients with neck complaints. *Spine*, *34*(4), E130–138. http://doi.org/10.1097/BRS.0b013e318195a28b

Humphry, R., Jewell, K., & Rosenberger, R. C. (1995). Development of in-hand manipulation and relationship with activities. *The American Journal of Occupational Therapy: Official Publication of the American Occupational Therapy Association*, *49*(8), 763–771.

Hung, M., Clegg, D. O., Greene, T., & Saltzman, C. L. (2011). Evaluation of the PROMIS physical function item bank in orthopaedic patients. *Journal of Orthopaedic Research: Official Publication of the Orthopaedic Research Society*, *29*(6), 947–953. http://doi.org/10.1002/jor.21308

Hwang, E. J., Hauschild, M., Wilke, M., & Andersen, R. A. (2012). Inactivation of the parietal reach region causes optic ataxia, impairing reaches but not saccades. *Neuron*, *76*(5), 1021–1029. http://doi.org/10.1016/j.neuron.2012.10.030

Ietswaart, M., Johnston, M., Dijkerman, H. C., Joice, S., Scott, C. L., MacWalter, R. S., & Hamilton, S. J. C. (2011). Mental practice with motor imagery in stroke recovery: randomized controlled trial of efficacy. *Brain: A Journal of Neurology*, *134*(Pt 5), 1373–1386. http://doi.org/10.1093/brain/awr077

Imaeda, T., Uchiyama, S., Wada, T., Okinaga, S., Sawaizumi, T., Omokawa, S., … Clinical Outcomes Committee of the Japanese Orthopaedic Association and the Functional Evaluation Committee of the Japanese Society for Surgery of the Hand. (2010). Reliability, validity, and responsiveness of the Japanese version of the Patient-Rated Wrist Evaluation. *Journal of Orthopaedic Science: Official Journal of the Japanese Orthopaedic Association*, *15*(4), 509–517. http://doi.org/10.1007/s00776-010-1477-x

Jacob-Lloyd, H., & Dunn, O. (2005). Effective measurement of the functional progress of stroke clients. *The British Journal of Occupational Therapy*, *68*(6), 253–59.

Jebsen, R. H., Taylor, N., Trieschmann, R. B., Trotter, M. J., & Howard, L. A. (1969). An objective and standardized test of hand function. *Archives of Physical Medicine and Rehabilitation*, *50*(6), 311–319.

Joint Committee on Testing Practices. (2004). Code of fair testing practices in education. National Council on Measurement in Education. Retrieved from http://www.apa.org/science/programs/testing/fair-testing.pdf

Jones, L. A., & Lederman, S. J. (2006). *Human hand function*. Oxford; New York: Oxford University Press.

Kandel, E. R., Schwartz, J. H., Jessell, T. M., Siegelbaum, S. A., & Hudspeth, A. J. (Eds.). (2012). *Principles of Neural Science, Fifth Edition (Principles of Neural Science* (5th ed.). McGraw-Hill Professional.

Kellor, M., Frost, J., Silberberg, N., Iversen, I., & Cummings, R. (1971). Hand strength and dexterity. *The American Journal of Occupational Therapy: Official Publication of the American Occupational Therapy Association*, *25*(2), 77–83.

Kennedy, C. A., Beaton, D. E., Smith, P., Van Eerd, D., Tang, K., Inrig, T., … Couban, R. (2013). Measurement properties of the QuickDASH (disabilities of the arm, shoulder and hand) outcome measure and cross-cultural adaptations of the QuickDASH: a systematic review. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, *22*(9), 2509–2547. http://doi.org/10.1007/s11136-013-0362-4

Ketterlin-Geller, L. R., Yovanoff, P., Jung, E., Liu, K., & Geller, J. (2013). Construct Definition Using Cognitively Based Evidence: A Framework for Practice. *Educational Assessment*, *18*(2), 122–146. http://doi.org/10.1080/10627197.2013.790207

Knutson, J. S., Harley, M. Y., Hisel, T. Z., Hogan, S. D., Maloney, M. M., & Chae, J. (2012). Contralaterally controlled functional electrical stimulation for upper extremity hemiplegia: an early-phase randomized clinical trial in subacute stroke patients. *Neurorehabilitation and Neural Repair*, *26*(3), 239–246. http://doi.org/10.1177/1545968311419301

Kopp, B., Kunkel, A., Flor, H., Platz, T., Rose, U., Mauritz, K. H., … Taub, E. (1997a). The Arm Motor Ability Test: reliability, validity, and sensitivity to change of an instrument for assessing disabilities in activities of daily living. *Archives of Physical Medicine and Rehabilitation*, *78*(6), 615–620.

Kopp, B., Kunkel, A., Flor, H., Platz, T., Rose, U., Mauritz, K. H., … Taub, E. (1997b). The Arm Motor Ability Test: reliability, validity, and sensitivity to change of an instrument for assessing disabilities in activities of daily living. *Archives of Physical Medicine and Rehabilitation*, *78*(6), 615–620.

Kraft, G. H., Amtmann, D., Bennett, S. E., Finlayson, M., Sutliff, M. H., Tullman, M., … Rabinowicz, A. L. (2014). Assessment of upper extremity function in multiple sclerosis: review and opinion. *Postgraduate Medicine*, *126*(5), 102–108. http://doi.org/10.3810/pgm.2014.09.2803

Krebs, H. I., Mernoff, S., Fasoli, S. E., Hughes, R., Stein, J., & Hogan, N. (2008). A comparison of functional and impairment-based robotic training in severe to moderate chronic stroke: a pilot study. *NeuroRehabilitation*, *23*(1), 81–87.

Krosnick, J., Holbrook, A., Berent, M., Carson, R., Hanemann, W., & Kopp, R. (2002). The impact of “no opinion” response options on data quality: Non-attitude reduction or an invitation to satisfice? *Public Opinion Quarterly*, *66*(3), 371–403.

Kunkel, A., Kopp, B., Müller, G., Villringer, K., Villringer, A., Taub, E., & Flor, H. (1999). Constraint-induced movement therapy for motor recovery in chronic stroke patients. *Archives of Physical Medicine and Rehabilitation*, *80*(6), 624–628.

Kurtaiş, Y., Küçükdeveci, A., Elhan, A., Yilmaz, A., Kalli, T., Tur, B. S., & Tennant, A. (2009). Psychometric properties of the Rivermead Motor Assessment: its utility in stroke. *Journal of Rehabilitation Medicine*, *41*(13), 1055–1061. http://doi.org/10.2340/16501977-0463

Lamers, I., Kerkhofs, L., Raats, J., Kos, D., Van Wijmeersch, B., & Feys, P. (2013). Perceived and actual arm performance in multiple sclerosis: relationship with clinical tests according to hand dominance. *Multiple Sclerosis (Houndmills, Basingstoke, England)*, *19*(10), 1341–1348. http://doi.org/10.1177/1352458513475832

Lamers, I., Timmermans, A., Kerkhofs, L., Severijns, D., Van Wijmeersch, B., & Feys, P. (2013). Self-reported use of the upper limbs related to clinical tests in persons with multiple sclerosis. *Disability and Rehabilitation*, *35*(23), 2016–2020. http://doi.org/10.3109/09638288.2013.771703

Lang, C. E., Edwards, D. F., Birkenmeier, R. L., & Dromerick, A. W. (2008). Estimating minimal clinically important differences of upper-extremity measures early after stroke. *Archives of Physical Medicine and Rehabilitation*, *89*(9), 1693–1700. http://doi.org/10.1016/j.apmr.2008.02.022

Lang, C. E., Wagner, J. M., Dromerick, A. W., & Edwards, D. F. (2006). Measurement of upper-extremity function early after stroke: properties of the action research arm test. *Archives of Physical Medicine and Rehabilitation*, *87*(12), 1605–1610. http://doi.org/10.1016/j.apmr.2006.09.003

Lannin, N., McCluskey, A., Cusick, A., Ashford, S., & Ross, L. (2010). Measuring function in everyday life: enhancing the Disabilities of the Arm Shoulder Hand questionnaire for use post-stroke. Presented at the World Federation of Occcupational Therapy, Santiago, Chile.

Lawshe, C. (1975). A quantitative approach to content validity. *Personnel Psychology*, *28*, 563–575. http://doi.org/10.1111/j.1744-6570.1975.tb01393.x

Lehman, L. A., Woodbury, M., Shechtman, O., Wang, Y. C., Pomeranz, J., Gray, D. B., & Velozo, C. A. (2011). Development of an item bank for a computerised adaptive test of upper-extremity function. *Disability and Rehabilitation*, *33*(21-22), 2092–2104. http://doi.org/10.3109/09638288.2011.560336

Lemmens, R. J. M., Timmermans, A. A. A., Janssen-Potten, Y. J. M., Smeets, R. J. E. M., & Seelen, H. A. M. (2012). Valid and reliable instruments for arm-hand assessment at ICF activity level in persons with hemiplegia: a systematic review. *BMC Neurology*, *12*, 21. http://doi.org/10.1186/1471-2377-12-21

Levy, R., Ruland, S., Weinand, M., Lowry, D., Dafer, R., & Bakay, R. (2008). Cortical stimulation for the rehabilitation of patients with hemiparetic stroke: a multicenter feasibility study of safety and efficacy. *Journal of Neurosurgery*, *108*(4), 707–714. http://doi.org/10.3171/JNS/2008/108/4/0707

Lincoln, N., & Leadbitter, D. (1979). Assessment of motor function in stroke patients. *Physiotherapy*, *65*(2), 48–51.

Lin, J.H., Hsueh, I.P., Sheu, C.F., & Hsieh, C.L. (2004). Psychometric properties of the sensory scale of the Fugl-Meyer Assessment in stroke patients. *Clinical Rehabilitation*, *18*(4), 391–397.

Lin, J.H., Hsu, M.J., Sheu, C.F., Wu, T.S., Lin, R.T., Chen, C.H., & Hsieh, C.L. (2009). Psychometric comparisons of 4 measures for assessing upper-extremity function in people with stroke. *Physical Therapy*, *89*(8), 840–850. http://doi.org/10.2522/ptj.20080285

Lin, K.C., Fu, T., Wu, C.Y., Hsieh, Y.W., Chen, C.L., & Lee, P.C. (2010). Psychometric comparisons of the Stroke Impact Scale 3.0 and Stroke-Specific Quality of Life Scale. *Quality of Life Research: An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, *19*(3), 435–443. http://doi.org/10.1007/s11136-010-9597-5

Lin, K., Chang, Y., Wu, C., & Chen, Y. (2009). Effects of constraint-induced therapy versus bilateral arm training on motor performance, daily functions, and quality of life in stroke survivors. *Neurorehabilitation and Neural Repair*, *23*(5), 441–448. http://doi.org/10.1177/1545968308328719

Lin, K., Chuang, L., Wu, C., Hsieh, Y., & Chang, W. (2010). Responsiveness and validity of three dexterous function measures in stroke rehabilitation. *Journal of Rehabilitation Research and Development*, *47*(6), 563–571.

Lin, K.C., Huang, Y.H., Hsieh, Y.W., & Wu, C.Y. (2009). Potential predictors of motor and functional outcomes after distributed constraint-induced therapy for patients with stroke. *Neurorehabilitation and Neural Repair*, *23*(4), 336–342. http://doi.org/10.1177/1545968308321773

Lin, K.C., Wu, C.Y., Wei, T.H., Lee, C.Y., & Liu, J. S. (2007). Effects of modified constraint-induced movement therapy on reach-to-grasp movements and functional performance after chronic stroke: a randomized controlled study. *Clinical Rehabilitation*, *21*(12), 1075–1086. http://doi.org/10.1177/0269215507079843

Lin, K., Fu, T., Wu, C., Wang, Y., Liu, J., Hsieh, C., & Lin, S. (2010). Minimal detectable change and clinically important difference of the Stroke Impact Scale in stroke patients. *Neurorehabilitation and Neural Repair*, *24*(5), 486–492. http://doi.org/10.1177/1545968309356295

Lin, K., Wu, C., Liu, J., Chen, Y., & Hsu, C. (2009). Constraint-induced therapy versus dose-matched control intervention to improve motor ability, basic/extended daily functions, and quality of life in stroke. *Neurorehabilitation and Neural Repair*, *23*(2), 160–165. http://doi.org/10.1177/1545968308320642

Livingston, S., & Zieky, M. (1982). Passing Scores: A Manual for Setting Standards of Performance on Educational and Occupational Tests. Educational Testing Service, Princeton, NJ. Retrieved from https://www.ets.org/Media/Research/pdf/passing_scores.pdf

Lundy-Ekman, L. (2012). *Neuroscience: Fundamentals for Rehabilitation, 4e* (4th ed.). Saunders.

Lyle, R. C. (1981). A performance test for assessment of upper limb function in physical rehabilitation treatment and research: *International Journal of Rehabilitation Research*, *4*(4), 483–492. http://doi.org/10.1097/00004356-198112000-00001

MacDermid, J. C., Wessel, J., Humphrey, R., Ross, D., & Roth, J. H. (2007). Validity of self-report measures of pain and disability for persons who have undergone arthroplasty for osteoarthritis of the carpometacarpal joint of the hand. *Osteoarthritis and Cartilage /*

OARS, Osteoarthritis Research Society, *15*(5), 524–530. http://doi.org/10.1016/j.joca.2006.10.018

Malouin, F., Pichard, L., Bonneau, C., Durand, A., & Corriveau, D. (1994). Evaluating motor recovery early after stroke: comparison of the Fugl-Meyer Assessment and the Motor Assessment Scale. *Archives of Physical Medicine and Rehabilitation*, *75*(11), 1206–1212.

Mao, H. F., Hsueh, I. P., Tang, P. F., Sheu, C. F., & Hsieh, C. L. (2002). Analysis and comparison of the psychometric properties of three balance measures for stroke patients. *Stroke; a Journal of Cerebral Circulation*, *33*(4), 1022–1027.

Marino, R. J., Kern, S. B., Leiby, B., Schmidt-Read, M., & Mulcahey, M. J. (2015). Reliability and validity of the capabilities of upper extremity test (CUE-T) in subjects with chronic spinal cord injury. *The Journal of Spinal Cord Medicine*, *38*(4), 498–504. http://doi.org/10.1179/2045772314Y.0000000272

Marshall, M. (1996). Sampling for qualitative research. *Family Practice*, *13*, 522–525.

Masiero, S., Celia, A., Rosati, G., & Armani, M. (2007). Robotic-assisted rehabilitation of the upper limb after acute stroke. *Archives of Physical Medicine and Rehabilitation*, *88*(2), 142–149. http://doi.org/10.1016/j.apmr.2006.10.032

Mathiowetz, V., Kashman, N., Volland, G., Weber, K., Dowe, M., & Rogers, S. (1985). Grip and pinch strength: normative data for adults. *Archives of Physical Medicine and Rehabilitation*, *66*(2), 69–74.

McDonnell, M. N., Hillier, S. L., Ridding, M. C., & Miles, T. S. (2006). Impairments in precision grip correlate with functional measures in adult hemiplegia. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *117*(7), 1474–1480. http://doi.org/10.1016/j.clinph.2006.02.027

McDowell, I. (2006). *Measuring Health:A Guide to Rating Scales and Questionnaires* (3rd ed.). Oxford University Press, USA.

Mehta, S., Macdermid, J. C., Carlesso, L. C., & McPhee, C. (2010). Concurrent validation of the DASH and the QuickDASH in comparison to neck-specific scales in patients with neck pain. *Spine*, *35*(24), 2150–2156. http://doi.org/10.1097/BRS.0b013e3181c85151

Mesquita, R., Janssen, D. J. A., Wouters, E. F. M., Schols, J. M. G. A., Pitta, F., & Spruit, M. A. (2013). Within-Day Test-Retest Reliability of the Timed Up & Go Test in Patients With Advanced Chronic Organ Failure. *Archives of Physical Medicine and Rehabilitation*. http://doi.org/10.1016/j.apmr.2013.03.024

Messick, S. (1989). Meaning and Values in Test Validation: The Science and Ethics of Assessment. *Educational Researcher*, *18*(2), 5–11.

Messick, S. (1994). The Interplay of Evidence and Consequences in the Validation of Performance Assessments. *Educational Researcher*, *23*(2), 13–23.

Metcalf, C., Adams, J., Burridge, J., Yule, V., & Chappell, P. (2007). A review of clinical upper limb assessments within the framework of the WHO ICF. *Musculoskeletal Care*, *5*(3), 160–173. http://doi.org/10.1002/msc.108

Milot, M.H., Spencer, S. J., Chan, V., Allington, J. P., Klein, J., Chou, C., … Cramer, S. C. (2014). Corticospinal excitability as a predictor of functional gains at the affected upper limb following robotic training in chronic stroke survivors. *Neurorehabilitation and Neural Repair*, *28*(9), 819–827. http://doi.org/10.1177/1545968314527351

Mislevy, R. J. (2007). Validity by Design. *Educational Researcher*, *36*(8), 463–469.

Mislevy, R. J. (2011, July). Evidence-centered design for simulation-based assessment. National Center for Rearch on Evaluation, Standards and Student Testing, Center for the study of evaluation, Graduate School of Education and Information Studies, UCLA, Los Angeles.

Mislevy, R. J., Almond, R. G., & Lukas, J. F. (2003, July). A Brief Introduction to Evidence-centered Design. Educational Testing Service, Princeton, NJ.

Miyamoto, S., Kondo, T., Suzukamo, Y., Michimata, A., & Izumi, S.I. (2009). Reliability and validity of the Manual Function Test in patients with stroke. *American Journal of Physical Medicine & Rehabilitation / Association of Academic Physiatrists*, *88*(3), 247–255. http://doi.org/10.1097/PHM.0b013e3181951133

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Medicine*, *6*(7), e1000097. http://doi.org/10.1371/journal.pmed.1000097

Morgan, S. J., Amtmann, D., Abrahamson, D. C., Kajlich, A. J., & Hafner, B. J. (2014). Use of cognitive interviews in the development of the PLUS-M item bank. *Quality of Life Research : An International Journal of Quality of Life Aspects of Treatment, Care and Rehabilitation*, *23*(6), 1767–1775. http://doi.org/10.1007/s11136-013-0618-z

Morris, D., Uswatte, G., Crago, J., Cook, E. W., 3rd, & Taub, E. (2001). The reliability of the wolf motor function test for assessing upper extremity function after stroke. *Archives of Physical Medicine and Rehabilitation*, *82*(6), 750–755. http://doi.org/10.1053/apmr.2001.23183

Morris, J. H., & Van Wijck, F. (2012). Responses of the less affected arm to bilateral upper limb task training in early rehabilitation after stroke: a randomized controlled trial. *Archives of*

Physical Medicine and Rehabilitation, *93*(7), 1129–1137. http://doi.org/10.1016/j.apmr.2012.02.025

Morris, J. H., van Wijck, F., Joice, S., Ogston, S. A., Cole, I., & MacWalter, R. S. (2008). A comparison of bilateral and unilateral upper-limb task training in early poststroke rehabilitation: a randomized controlled trial. *Archives of Physical Medicine and Rehabilitation*, *89*(7), 1237–1245. http://doi.org/10.1016/j.apmr.2007.11.039

Mulcahey, M., Smith, B., & Betz, R. (2004). Psychometric Rigor of the Grasp and Release Test for Measuring Funcitonal Limitation of Persons with Tetraplegia: A Preliminary analysis. *The Journal of Spinal Cord Medicine*, *27*(1), 41–46.

Napier, J., & Tuttle, R. H. (1993). *Hands* (Rev Sub edition). Princeton, N.J: Princeton University Press.

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., … Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, *53*(4), 695–699.

Nesselroade, J. R., Gerstorf, D., Hardy, S. A., & Ram, N. (2007). Idiographic Filters for Psychological Constructs. *Measurement: Interdisciplinary Research and Perspectives*, *5*(4), 217–235.

Netemeyer, R., Bearden, W., & Sharma, S. (2003). *Scaling Procedures: Issues and Applications* (1 edition). Thousand Oaks, CA: SAGE Publications, Inc.

NIH Consensus Development Panel on Rehabilitation of Persons With Traumatic Brain Injury. (1999). Rehabilitation of persons with traumatic brain injury. *JAMA*, *282*(10), 974–983. http://doi.org/10.1001/jama.282.10.974

Nijland, R., van Wegen, E., Verbunt, J., van Wijk, R., van Kordelaar, J., & Kwakkel, G. (2010). A comparison of two validated tests for upper limb function after stroke: The Wolf Motor Function Test and the Action Research Arm Test. *Journal of Rehabilitation Medicine*, *42*(7), 694–696. http://doi.org/10.2340/16501977-0560

Ohn, S. H., Yoo, W. K., Kim, D. Y., Ahn, S., Jung, B., Choi, I., … Jung, K. I. (2013). Measurement of synergy and spasticity during functional movement of the post-stoke hemiplegic upper limb. *Journal of Electromyography and Kinesiology: Official Journal of the International Society of Electrophysiological Kinesiology*, *23*(2), 501–507. http://doi.org/10.1016/j.jelekin.2012.10.001

Oxford Grice, K., Vogel, K. A., Le, V., Mitchell, A., Muniz, S., & Vollmer, M. A. (2003). Adult norms for a commercially available Nine Hole Peg Test for finger dexterity. *The American Journal of Occupational Therapy: Official Publication of the American Occupational Therapy Association*, *57*(5), 570–573.

Page, S. J., Hade, E., & Persch, A. (2015). Psychometrics of the wrist stability and hand mobility subscales of the Fugl-Meyer assessment in moderately impaired stroke. *Physical Therapy*, *95*(1), 103–108. http://doi.org/10.2522/ptj.20130235

Page, S. J., Hermann, V. H., Levine, P. G., Lewis, E., Stein, J., & DePeel, J. (2011). Portable neurorobotics for the severely affected arm in chronic stroke: a case study. *Journal of Neurologic Physical Therapy: JNPT*, *35*(1), 41–46. http://doi.org/10.1097/NPT.0b013e31820c4b00

Page, S. J., Levine, P., & Hade, E. (2012). Psychometric properties and administration of the wrist/hand subscales of the Fugl-Meyer Assessment in minimally impaired upper extremity hemiparesis in stroke. *Archives of Physical Medicine and Rehabilitation*, *93*(12), 2373–2376.e5. http://doi.org/10.1016/j.apmr.2012.06.017

Pang, M. Y., Harris, J. E., & Eng, J. J. (2006). A community-based upper-extremity group exercise program improves motor function and performance of functional activities in chronic stroke: a randomized controlled trial. *Archives of Physical Medicine and Rehabilitation*, *87*(1), 1–9. http://doi.org/10.1016/j.apmr.2005.08.113

Patient-Centered Outcomes Research Institute. (2012, June 4). Preliminary draft methodology report. Retrieved January 5, 2016, from http://www.pcori.org/assets/Preliminary-Draft-Methodology-Report.pdf

Pendleton, H., & Schultz-Krohn, W. (2006). *Pedretti's Occupational Therapy 6th* (6th edition). Mosby Co.

Penta, M. (2015). ABILHAND - Presentation - Rehab-Scales.org. Retrieved October 22, 2015, from http://www.rehab-scales.org/abilhand-presentation.html#development-of-the-abilhand-scale-using-the-rasch-measurement-model

Penta, M., Tesio, L., Arnould, C., Zancan, A., & Thonnard, J. L. (2001). The ABILHAND questionnaire as a measure of manual ability in chronic stroke patients: Rasch-based validation and relationship to upper limb impairment. *Stroke; a Journal of Cerebral Circulation*, *32*(7), 1627–1634.

Penta, M., Thonnard, J. L., & Tesio, L. (1998). ABILHAND: a Rasch-built measure of manual ability. *Archives of Physical Medicine and Rehabilitation*, *79*(9), 1038–1042.

Peters, M. J. H., van Nes, S. I., Vanhoutte, E. K., Bakkers, M., van Doorn, P. A., Merkies, I. S. J., … PeriNomS Study group. (2011). Revised normative values for grip strength with the

Jamar dynamometer. *Journal of the Peripheral Nervous System: JPNS*, *16*(1), 47–50. http://doi.org/10.1111/j.1529-8027.2011.00318.x

Pfeiffer, E. (1975). A short portable mental status questionnaire for the assessment of organic brain deficit in elderly patients. *Journal of the American Geriatrics Society*, *23*(10), 433–441.

Platz, T., Pinkowski, C., van Wijck, F., Kim, I. H., di Bella, P., & Johnson, G. (2005). Reliability and validity of arm function assessment with standardized guidelines for the Fugl-Meyer Test, Action Research Arm Test and Box and Block Test: a multicentre study. *Clinical Rehabilitation*, *19*(4), 404–411.

Polman, C. H., & Rudick, R. A. (2010). The multiple sclerosis functional composite: a clinically meaningful measure of disability. *Neurology*, *74 Suppl 3*, S8–15. http://doi.org/10.1212/WNL.0b013e3181dbb571

Pomeroy, V. M., Ward, N. S., Johansen-Berg, H., van Vliet, P., Burridge, J., Hunter, S. M., … FAST INdICATE Investigators. (2014). FAST INdiCATE Trial protocol. Clinical efficacy of functional strength training for upper limb motor recovery early after stroke: neural correlates and prognostic indicators. *International Journal of Stroke: Official Journal of the International Stroke Society*, *9*(2), 240–245. http://doi.org/10.1111/ijs.12179

Quatrano, L. A., & Cruz, T. H. (2011). Future of outcomes measurement: impact on research in medical rehabilitation and neurologic populations. *Archives of Physical Medicine and Rehabilitation*, *92*(10 Suppl), S7–11. http://doi.org/10.1016/j.apmr.2010.08.032

Rabadi, M. H., & Rabadi, F. M. (2006). Comparison of the action research arm test and the Fugl-Meyer assessment as measures of upper-extremity motor weakness after stroke. *Archives* *of Physical Medicine and Rehabilitation*, *87*(7), 962–966. http://doi.org/10.1016/j.apmr.2006.02.036

Radomski, M., & Trombly, C. A. (Eds.). (2007). *Occupational Therapy for Physical Dysfunction* (Sixth). Philadelphia: Lippincott Williams & Wilkins.

Rallon, C. R., & Chen, C. C. (2008). Relationship between performance-based and self-reported assessment of hand function. *The American Journal of Occupational Therapy: Official Publication of the American Occupational Therapy Association*, *62*(5), 574–579.

Rapin, I., Tourk, L. M., & Costa, L. D. (1966). Evaluation of the Purdue Pegboard as a screening test for brain damage. *Developmental Medicine and Child Neurology*, *8*(1), 45–54.

Reinkensmeyer, D. J., Wolbrecht, E. T., Chan, V., Chou, C., Cramer, S. C., & Bobrow, J. E. (2012). Comparison of three-dimensional, assist-as-needed robotic arm/hand movement training provided with Pneu-WREX to conventional tabletop therapy after chronic stroke. *American Journal of Physical Medicine & Rehabilitation / Association of Academic Physiatrists*, *91*(11 Suppl 3), S232–241. http://doi.org/10.1097/PHM.0b013e31826bce79

Rosati, G., Gallina, P., & Masiero, S. (2007). Design, implementation and clinical tests of a wire-based robot for neurorehabilitation. *IEEE Transactions on Neural Systems and Rehabilitation Engineering: A Publication of the IEEE Engineering in Medicine and Biology Society*, *15*(4), 560–569. http://doi.org/10.1109/TNSRE.2007.908560

Rosewilliam, S., Malhotra, S., Roffe, C., Jones, P., & Pandyan, A. D. (2012). Can surface neuromuscular electrical stimulation of the wrist and hand combined with routine therapy facilitate recovery of arm function in patients with stroke? *Archives of Physical Medicine and Rehabilitation*, *93*(10), 1715–1721.e1. http://doi.org/10.1016/j.apmr.2012.05.017

Sabari, J. S., Lim, A. L., Velozo, C. A., Lehman, L., Kieran, O., & Lai, J. S. (2005). Assessing arm and hand function after stroke: a validity test of the hierarchical scoring system used in the motor assessment scale for stroke. *Archives of Physical Medicine and Rehabilitation*, *86*(8), 1609–1615. http://doi.org/10.1016/j.apmr.2004.12.028

Sackley, C. M., & Lincoln, N. B. (1990). The verbal administration of the gross function scale of the Rivermead Motor Assessment. *Clinical Rehabilitation*, *4*(4), 301–303. http://doi.org/10.1177/026921559000400409

Sadler, J. Z. (2003). *Descriptions and Prescriptions: Values, Mental Disorders, and the DSMs*. Baltimore, MD. The John Hopkins University Press.

Sawaki, L., Butler, A. J., Leng, X., Wassenaar, P. A., Mohammad, Y. M., Blanton, S., … Wittenberg, G. F. (2008). Constraint-induced movement therapy results in increased motor map area in subjects 3 to 9 months after stroke. *Neurorehabilitation and Neural Repair*, *22*(5), 505–513. http://doi.org/10.1177/1545968308317531

Schuster-Amft, C., Eng, K., Lehmann, I., Schmid, L., Kobashi, N., Thaler, I., … Kiper, D. (2014). Using mixed methods to evaluate efficacy and user expectations of a virtual reality-based training system for upper-limb recovery in patients after stroke: a study protocol for a randomised controlled trial. *Trials*, *15*, 350. http://doi.org/10.1186/1745-6215-15-350

Sears, E. D., & Chung, K. C. (2010). Validity and responsiveness of the Jebsen-Taylor Hand Function Test. *The Journal of Hand Surgery*, *35*(1), 30–37. http://doi.org/10.1016/j.jhsa.2009.09.008

Seo, N. J., Fischer, H. W., Bogey, R. A., Rymer, W. Z., & Kamper, D. G. (2011). Use of visual force feedback to improve digit force direction during pinch grip in persons with stroke: a pilot study. *Archives of Physical Medicine and Rehabilitation*, *92*(1), 24–30. http://doi.org/10.1016/j.apmr.2010.08.016

Shaw, L., Rodgers, H., Price, C., van Wijck, F., Shackley, P., Steen, N., … BoTULS investigators. (2010). BoTULS: a multicentre randomised controlled trial to evaluate the clinical effectiveness and cost-effectiveness of treating upper limb spasticity due to stroke with botulinum toxin type A. *Health Technology Assessment (Winchester, England)*, *14*(26), 1–113, iii–iv. http://doi.org/10.3310/hta14260

Shaw, S. E., Morris, D. M., Uswatte, G., McKay, S., Meythaler, J. M., & Taub, E. (2005). Constraint-induced movement therapy for recovery of upper-limb function following traumatic brain injury. *Journal of Rehabilitation Research and Development*, *42*(6), 769–778.

Shelton, F. D., Volpe, B. T., & Reding, M. (2001). Motor impairment as a predictor of functional recovery and guide to rehabilitation treatment after stroke. *Neurorehabilitation and Neural Repair*, *15*(3), 229–237.

Shumway-Cook, A., & Woollacott, M. (2011). *Motor Control: Translating Research into Clinical Practice* (Fourth, North American Edition). Philadelphia: LWW.

Siebers, A., Oberg, U., & Skargren, E. (2010). The effect of modified constraint-induced movement therapy on spasticity and motor function of the affected arm in patients with chronic stroke. *Physiotherapy Canada. Physiothérapie Canada*, *62*(4), 388–396. http://doi.org/10.3138/physio.62.4.388

Simondson, J. A., Goldie, P., & Greenwood, K. M. (2003). The Mobility Scale for Acute Stroke Patients: concurrent validity. *Clinical Rehabilitation*, *17*(5), 558–564. http://doi.org/10.1191/0269215503cr650oa

Simone, A., Rota, V., Tesio, L., & Perucca, L. (2011). Generic ABILHAND questionnaire can measure manual ability across a variety of motor impairments. *International Journal of Rehabilitation Research. Internationale Zeitschrift Für Rehabilitationsforschung. Revue Internationale De Recherches De Réadaptation*, *34*(2), 131–140. http://doi.org/10.1097/MRR.0b013e328343d4d3

Skirven, T., Osterman, L., Fedorczyk, J., & Amadio, P. (2011). *Rehabilitation of the hand and upper extremity*. Philadelphia, PA: Elsevier Mosby. Retrieved from http://www.mdconsult.com/public/book/view?title=Skirven:+Rehabilitation+of+the+Hand+and+Upper+Extremity

Slobogean, G. P., Noonan, V. K., Famuyide, A., & O'Brien, P. J. (2011). Does objective shoulder impairment explain patient-reported functional outcome? A study of proximal humerus fractures. *Journal of Shoulder and Elbow Surgery / American Shoulder and Elbow Surgeons*, *20*(2), 267–272. http://doi.org/10.1016/j.jse.2010.06.005

Sollerman, C., & Ejeskär, A. (1995). Sollerman hand function test. A standardised method and its use in tetraplegic patients. *Scandinavian Journal of Plastic and Reconstructive Surgery and Hand Surgery / Nordisk Plastikkirurgisk Forening [and] Nordisk Klubb for Handkirurgi*, *29*(2), 167–176.

Soyuer, F., & Soyuer, A. (2005). Ischemic stroke: Motor impairment and disability with relation to age and lesion location. *Journal of Neurological Sciences (Turkish)*, *22*(1), 043–049.

Stamm, T. A., Cieza, A., Machold, K. P., Smolen, J. S., & Stucki, G. (2004). Content comparison of occupation-based instruments in adult rheumatology and musculoskeletal rehabilitation based on the International Classification of Functioning, Disability and Health. *Arthritis and Rheumatism*, *51*(6), 917–924. http://doi.org/10.1002/art.20842

Stern, E. B. (1992). Stability of the Jebsen-Taylor Hand Function Test across three test sessions. *The American Journal of Occupational Therapy: Official Publication of the American Occupational Therapy Association*, *46*(7), 647–649.

Sullivan, J. E., Hurley, D., & Hedman, L. D. (2012). Afferent stimulation provided by glove electrode during task-specific arm exercise following stroke. *Clinical Rehabilitation*, *26*(11), 1010–1020. http://doi.org/10.1177/0269215512442915

Sunderland, A., Tinson, D., Bradley, L., & Hewer, R. L. (1989). Arm function after stroke. An evaluation of grip strength as a measure of recovery and a prognostic indicator. *Journal of Neurology, Neurosurgery, and Psychiatry*, *52*(11), 1267–1272.

Suzuki, M., Yamada, S., Omori, M., Hatakeyama, M., Sugimura, Y., Matsushita, K., & Tagawa, Y. (2008). Development of the upper-body dressing scale for a buttoned shirt: a preliminary correlational study. *American Journal of Physical Medicine & Rehabilitation / Association of Academic Physiatrists*, *87*(9), 740–749. http://doi.org/10.1097/PHM.0b013e31818378b0

Taub, E., Uswatte, G., Bowman, M., Mark, V., Delgado, A., Bryson, C., … Bishop-McKay, S. (2013). Constraint-induced movement therapy combined with conventional neurorehabilitation techniques in chronic stroke patients with plegic hands: a case series. *Archives of Physical Medicine and Rehabilitation*, *94*(1), 86–94. http://doi.org/10.1016/j.apmr.2012.07.029

Timmermans, A. A. A., Lemmens, R. J. M., Monfrance, M., Geers, R. P. J., Bakx, W., Smeets, R. J. E. M., & Seelen, H. A. M. (2014). Effects of task-oriented robot training on arm function, activity, and quality of life in chronic stroke patients: a randomized controlled trial. *Journal of Neuroengineering and Rehabilitation*, *11*, 45. http://doi.org/10.1186/1743-0003-11-45

Turk, R., Burridge, J. H., Davis, R., Cosendai, G., Sparrow, O., Roberts, H. C., … Schulman, J. (2008). Therapeutic effectiveness of electric stimulation of the upper-limb poststroke using implanted microstimulators. *Archives of Physical Medicine and Rehabilitation*, *89*(10), 1913–1922. http://doi.org/10.1016/j.apmr.2008.01.030

Tyser, A. R., Beckmann, J., Franklin, J. D., Cheng, C., Hon, S. D., Wang, A., & Hung, M. (2014). Evaluation of the PROMIS physical function computer adaptive test in the upper extremity. *The Journal of Hand Surgery*, *39*(10), 2047–2051.e4. http://doi.org/10.1016/j.jhsa.2014.06.130

Tyson, S. F., & DeSouza, L. H. (2004). Reliability and validity of functional balance tests post stroke. *Clinical Rehabilitation*, *18*(8), 916–923.

Tyson, & Turner. (1999). The process of stroke rehabilitation: what happens and why. *Clinical Rehabilitation*, *13*(4), 322–332.

Tyson, & Turner. (2000). Discharge and follow-up for people with stroke: what happens and why. *Clinical Rehabilitation*, *14*(4), 381–392.

Urbin, M. A., Waddell, K. J., & Lang, C. E. (2015). Acceleration metrics are responsive to change in upper extremity function of stroke survivors. *Archives of Physical Medicine and Rehabilitation*, *96*(5), 854–861. http://doi.org/10.1016/j.apmr.2014.11.018

U.S. Department of Health and Human Services FDA Center for Drug Evaluation and Research, U.S. Department of Health and Human Services FDA Center for Biologics Evaluation and Research, & U.S. Department of Health and Human Services FDA Center for Devices and Radiological Health. (2006). Guidance for industry: patient-reported outcome measures: use in medical product development to support labeling claims: draft guidance. *Health and Quality of Life Outcomes*, *4*, 79. http://doi.org/10.1186/1477-7525-4-79

Uswatte, G., Taub, E., Morris, D., Light, K., & Thompson, P. A. (2006). The Motor Activity Log-28: assessing daily use of the hemiparetic arm after stroke. *Neurology*, *67*(7), 1189–1194. http://doi.org/10.1212/01.wnl.0000238164.90657.c2

Uswatte, G., Taub, E., Morris, D., Vignolo, M., & McCulloch, K. (2005). Reliability and validity of the upper-extremity Motor Activity Log-14 for measuring real-world arm use. *Stroke; a Journal of Cerebral Circulation*, *36*(11), 2493–2496.

van Delden, A. L. E. Q., Peper, C. L. E., Beek, P. J., & Kwakkel, G. (2013). Match and mismatch between objective and subjective improvements in upper limb function after stroke. *Disability and Rehabilitation*, *35*(23), 1961–1967. http://doi.org/10.3109/09638288.2013.768303

van Delden, A. L. E. Q., Peper, C. L. E., Nienhuys, K. N., Zijp, N. I., Beek, P. J., & Kwakkel, G. (2013). Unilateral versus bilateral upper limb training after stroke: the Upper Limb Training After Stroke clinical trial. *Stroke; a Journal of Cerebral Circulation*, *44*(9), 2613–2616. http://doi.org/10.1161/STROKEAHA.113.001969

van der Lee, J. H., Beckerman, H., Knol, D. L., de Vet, H. C. W., & Bouter, L. M. (2004). Clinimetric properties of the motor activity log for the assessment of arm use in hemiparetic patients. *Stroke; a Journal of Cerebral Circulation*, *35*(6), 1410–1414. http://doi.org/10.1161/01.STR.0000126900.24964.7e van der Lee, J. H., Beckerman, H., Lankhorst, G. J., & Bouter, L. M. (2001). The responsiveness of the Action Research Arm test and the Fugl-Meyer Assessment scale in chronic stroke patients. *Journal of Rehabilitation Medicine*, *33*(3), 110–113.

van der Lee, J. H., Wagenaar, R. C., Lankhorst, G. J., Vogelaar, T. W., Devillé, W. L., & Bouter, L. M. (1999). Forced use of the upper extremity in chronic stroke patients: results from a single-blind randomized clinical trial. *Stroke; a Journal of Cerebral Circulation*, *30*(11), 2369–2375.

van de Ven-Stevens, L. A. W., Graff, M. J. L., Selles, R. W., Schreuders, T. A. R., van der Linde, H., Spauwen, P. H., & Geurts, A. C. H. (2016). Instruments for assessment of impairments and activity limitations in patients with hand conditions: A European Delphi study. *Journal of Rehabilitation Medicine*, *47*(10), 948–956. http://doi.org/10.2340/16501977-2015

Veehof, M. M., Sleegers, E. J. A., van Veldhoven, N. H. M. J., Schuurman, A. H., & van Meeteren, N. L. U. (2002). Psychometric qualities of the Dutch language version of the Disabilities of the Arm, Shoulder, and Hand questionnaire (DASH-DLV). *Journal of Hand Therapy: Official Journal of the American Society of Hand Therapists*, *15*(4), 347–354.

Velozo, C. A., Seel, R. T., Magasi, S., Heinemann, A. W., & Romero, S. (2012). Improving measurement methods in rehabilitation: core concepts and recommendations for scale development. *Archives of Physical Medicine and Rehabilitation*, *93*(8 Suppl), S154–163.

Velstra, I. M., Ballert, C. S., & Cieza, A. (2011). A systematic literature review of outcome measures for upper extremity function using the international classification of functioning, disability, and health as reference. *PM & R: The Journal of Injury, Function, and Rehabilitation*, *3*(9), 846–860. http://doi.org/10.1016/j.pmrj.2011.03.014

Venkataraman, S. T., & Iberall, T. (2012). *Dextrous Robot Hands*. Springer Science & Business Media.

Vos-Vromans, D. C. W. M., de Bie, R. A., Erdmann, P. G., & van Meeteren, N. L. U. (2005). The responsiveness of the ten-meter walking test and other measures in patients with hemiparesis in the acute phase. *Physiotherapy Theory and Practice*, *21*(3), 173–180.

Vuillermot, S., Pescatore, A., Holper, L., Kiper, D. C., & Eng, K. (2009). An extended drawing test for the assessment of arm and hand function with a performance invariant for healthy subjects. *Journal of Neuroscience Methods*, *177*(2), 452–460. http://doi.org/10.1016/j.jneumeth.2008.10.018

Wagner, J. M., Rhodes, J. A., & Patten, C. (2008). Reproducibility and minimal detectable change of three-dimensional kinematic analysis of reaching tasks in people with hemiparesis after stroke. *Physical Therapy*, *88*(5), 652–663. http://doi.org/10.2522/ptj.20070255

Wang, T., Lin, K., Wu, C., Chung, C., Pei, Y., & Teng, Y. (2011). Validity, responsiveness, and clinically important difference of the ABILHAND questionnaire in patients with stroke. *Archives of Physical Medicine and Rehabilitation*, *92*(7), 1086–1091.

Washington University in St. Louis. (2009). Men in Nursing, Occupational Therapy, & Physical Therapy. Bernard Becker Medical Library. Retrieved from http://beckerexhibits.wustl.edu/mowihsp/stats/men.htm

Weng, L. Y., Hsieh, C. L., Tung, K. Y., Wang, T. J., Ou, Y. C., Chen, L. R., … Liu, C. F. (2010). Excellent reliability of the Sollerman hand function test for patients with burned hands. *Journal of Burn Care & Research: Official Publication of the American Burn Association*, *31*(6), 904–910. http://doi.org/10.1097/BCR.0b013e3181f93583

WHO. (2001). *International Classification of Functioning, Disability and Health.* Geneva: World Health Organization.

Williams, G., Robertson, V., Greenwood, K., Goldie, P., & Morris, M. E. (2006). The concurrent validity and responsiveness of the high-level mobility assessment tool for measuring the mobility limitations of people with traumatic brain injury. *Archives of Physical Medicine and Rehabilitation*, *87*(3), 437–442. http://doi.org/10.1016/j.apmr.2005.10.028

Wilson, D., Baker, L., & Craddock, J. (1984). Functional Test for the Hemiparetic Upper Extremity. *American Journal of Occupational Therapy*, *38*(3), 159–164.

Winstein, C. J., Wolf, S. L., Dromerick, A. W., Lane, C. J., Nelsen, M. A., Lewthwaite, R., … ICARE Investigative Team. (2013). Interdisciplinary Comprehensive Arm Rehabilitation Evaluation (ICARE): a randomized controlled trial protocol. *BMC Neurology*, *13*, 5. http://doi.org/10.1186/1471-2377-13-5

Wolf, S. L., Catlin, P. A., Ellis, M., Archer, A. L., Morgan, B., & Piacentino, A. (2001). Assessing Wolf motor function test as outcome measure for research in patients after stroke. *Stroke; a Journal of Cerebral Circulation*, *32*(7), 1635–1639.

Wolf, S. L., Winstein, C. J., Miller, J. P., Taub, E., Uswatte, G., Morris, D., … EXCITE Investigators. (2006). Effect of constraint-induced movement therapy on upper extremity function 3 to 9 months after stroke: the EXCITE randomized clinical trial. *JAMA*, *296*(17), 2095–2104. http://doi.org/10.1001/jama.296.17.2095

Woodbury, M. L., Velozo, C. A., Richards, L. G., Duncan, P. W., Studenski, S., & Lai, S. M. (2008). Longitudinal stability of the Fugl-Meyer Assessment of the upper extremity.

Archives of Physical Medicine and Rehabilitation, *89*(8), 1563–1569. http://doi.org/10.1016/j.apmr.2007.12.041

Wood-Dauphinee, S. L., Williams, J. I., & Shapiro, S. H. (1990). Examining outcome measures in a clinical study of stroke. *Stroke; a Journal of Cerebral Circulation*, *21*(5), 731–739.

Wu, C. W., Seo, H. J., & Cohen, L. G. (2006). Influence of electric somatosensory stimulation on paretic-hand function in chronic stroke. *Archives of Physical Medicine and Rehabilitation*, *87*(3), 351–357. http://doi.org/10.1016/j.apmr.2005.11.019

Yozbatiran, N., Der-Yeghiaian, L., & Cramer, S. C. (2008). A standardized approach to performing the Action Research Arm Test. *Neurorehabilitation and Neural Repair*, *22*(1), 78–90. http://doi.org/10.1177/1545968307305353